



Analyse expérimentale et numérique des défaillances mécaniques locales induites dans les interconnexions par les tests paramétriques et les assemblages : optimisation des procédés et des architectures des plots de connexion

Romuald Roucou

► To cite this version:

Romuald Roucou. Analyse expérimentale et numérique des défaillances mécaniques locales induites dans les interconnexions par les tests paramétriques et les assemblages : optimisation des procédés et des architectures des plots de connexion. Autre. Ecole Nationale Supérieure des Mines de Saint-Etienne, 2010. Français. NNT : 2010EMSE0589 . tel-00669635

HAL Id: tel-00669635

<https://theses.hal.science/tel-00669635>

Submitted on 13 Feb 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

N° d'ordre : 2010 EMSE 0589

THÈSE

présentée par

Romuald ROUCOU

Pour obtenir le grade de
Docteur de l'École Nationale Supérieure des Mines de Saint-Étienne

Spécialité : Microélectronique

Analyse expérimentale et numérique des défaillances mécaniques locales
induites dans les interconnexions par les tests paramétriques et les
assemblages : Optimisation des procédés et des architectures des plots de
connexion

soutenue à Gardanne, le 9 décembre 2010

Membres du jury

Président :	François BAY	Directeur de Recherche Ecole des Mines, MINES ParisTech, CEMEF, Sophia Antipolis
Rapporteurs :	Philippe GOUDEAU	Directeur de Recherche CNRS, Institut Pprime, CNRS/Université de Poitiers/ENSMA, Poitiers
	François BAY	Directeur de Recherche Ecole des Mines, MINES ParisTech, CEMEF, Sophia Antipolis
Examineurs :	Sandrine BEC	Chargé de recherche CNRS, LTDS, Ecole Centrale de Lyon Ingénieur R&D, STMicroelectronics, Crolles
	Vincent FIORI	Ingénieur R&D, STMicroelectronics, Agrate, Italie
	Lucia ZULLINO	Professeur Ecole des Mines, Ecole Nationale Supérieure des Mines de Saint-Etienne, Gardanne
Co-encadrant :	Xavier BODDAERT	Professeur Ecole des Mines, Ecole Nationale Supérieure des Mines de Saint-Etienne, Gardanne
Directeur de thèse :	Karim INAL	Mines de Saint-Etienne, Gardanne

Spécialités doctorales :

SCIENCES ET GENIE DES MATERIAUX
 MECANIQUE ET INGENIERIE
 GENIE DES PROCEDES
 SCIENCES DE LA TERRE
 SCIENCES ET GENIE DE L'ENVIRONNEMENT
 MATHEMATIQUES APPLIQUEES
 INFORMATIQUE
 IMAGE, VISION, SIGNAL
 GENIE INDUSTRIEL
 MICROELECTRONIQUE

Responsables :

J. DRIVER Directeur de recherche – Centre SMS
 A. VAUTRIN Professeur – Centre SMS
 G. THOMAS Professeur – Centre SPIN
 B. GUY Maître de recherche – Centre SPIN
 J. BOURGOIS Professeur – Centre SITE
 E. TOUBOUL Ingénieur – Centre G2I
 O. BOISSIER Professeur – Centre G2I
 JC. PINOLI Professeur – Centre CIS
 P. BURLAT Professeur – Centre G2I
 Ph. COLLOT Professeur – Centre CMP

Enseignants-chercheurs et chercheurs autorisés à diriger des thèses de doctorat (titulaires d'un doctorat d'État ou d'une HDR)

AVRIL	Stéphane	MA	Mécanique & Ingénierie	CIS
BATTON-HUBERT	Mireille	MA	Sciences & Génie de l'Environnement	SITE
BENABEN	Patrick	PR 1	Sciences & Génie des Matériaux	CMP
BERNACHE-ASSOLANT	Didier	PR 0	Génie des Procédés	CIS
BIGOT	Jean-Pierre	MR	Génie des Procédés	SPIN
BILAL	Essaïd	DR	Sciences de la Terre	SPIN
BOISSIER	Olivier	PR 1	Informatique	G2I
BOUCHER	Xavier	MA	Génie Industriel	G2I
BOUDAREL	Marie-Reine	PR 2	Génie Industriel	DF
BOURGOIS	Jacques	PR 0	Sciences & Génie de l'Environnement	SITE
BRODHAG	Christian	DR	Sciences & Génie de l'Environnement	SITE
BURLAT	Patrick	PR 2	Génie industriel	G2I
COLLOT	Philippe	PR 1	Microélectronique	CMP
COURNIL	Michel	PR 0	Génie des Procédés	SPIN
DAUZERE-PERES	Stéphane	PR 1	Génie industriel	CMP
DARRIEULAT	Michel	IGM	Sciences & Génie des Matériaux	SMS
DECHOMETS	Roland	PR 1	Sciences & Génie de l'Environnement	SITE
DESRAYAUD	Christophe	MA	Mécanique & Ingénierie	SMS
DELAFOSSÉ	David	PR 1	Sciences & Génie des Matériaux	SMS
DOLGUI	Alexandre	PR 1	Génie Industriel	G2I
DRAPIER	Sylvain	PR 2	Mécanique & Ingénierie	SMS
DRIVER	Julian	DR 0	Sciences & Génie des Matériaux	SMS
FEILLET	Dominique	PR 2	Génie Industriel	CMP
FOREST	Bernard	PR 1	Sciences & Génie des Matériaux	CIS
FORMISYN	Pascal	PR 1	Sciences & Génie de l'Environnement	SITE
FORTUNIER	Roland	PR 1	Sciences & Génie des Matériaux	SMS
FRACZKIEWICZ	Anna	DR	Sciences & Génie des Matériaux	SMS
GARCIA	Daniel	MR	Génie des Procédés	SPIN
GIRARDOT	Jean-Jacques	MR	Informatique	G2I
GOEURLOT	Dominique	MR	Sciences & Génie des Matériaux	SMS
GRAILLOT	Didier	DR	Sciences & Génie de l'Environnement	SITE
GROSSEAU	Philippe	MR	Génie des Procédés	SPIN
GRUY	Frédéric	MR	Génie des Procédés	SPIN
GUY	Bernard	MR	Sciences de la Terre	SPIN
GUYONNET	René	DR	Génie des Procédés	SPIN
HERRI	Jean-Michel	PR 2	Génie des Procédés	SPIN
INAL	Karim	PR 2	Microélectronique	CMP
KLÖCKER	Helmut	DR	Sciences & Génie des Matériaux	SMS
LAFOREST	Valérie	CR	Sciences & Génie de l'Environnement	SITE
LERICHE	Rodolphe	CR CNRS	Mécanique et Ingénierie	SMS
LI	Jean-Michel	EC (CCI MP)	Microélectronique	CMP
LONDICHE	Henry	MR	Sciences & Génie de l'Environnement	SITE
MALLIARAS	George Grégory	PR 1	Microélectronique	CMP
MOLIMARD	Jérôme	MA	Mécanique et Ingénierie	SMS
MONTHEILLET	Frank	DR 1 CNRS	Sciences & Génie des Matériaux	SMS
PERIER-CAMBY	Laurent	PR 2	Génie des Procédés	SPIN
PIJOLAT	Christophe	PR 1	Génie des Procédés	SPIN
PIJOLAT	Michèle	PR 1	Génie des Procédés	SPIN
PINOLI	Jean-Charles	PR 0	Image, Vision, Signal	CIS
STOLARZ	Jacques	CR	Sciences & Génie des Matériaux	SMS
SZAFNICKI	Konrad	MR	Sciences & Génie de l'Environnement	SITE
THOMAS	Gérard	PR 0	Génie des Procédés	SPIN
TRIA	Assia		Microélectronique	CMP
VALDIVIESO	François	MA	Sciences & Génie des Matériaux	SMS
VAUTRIN	Alain	PR 0	Mécanique & Ingénierie	SMS
VIRICELLE	Jean-Paul	MR	Génie des procédés	SPIN
WOLSKI	Krzysztof	DR	Sciences & Génie des Matériaux	SMS
XIE	Xiaolan	PR 1	Génie industriel	CIS

Glossaire :

PR 0 Professeur classe exceptionnelle
 PR 1 Professeur 1^{ère} catégorie
 PR 2 Professeur 2^{ème} catégorie
 MA(MDC) Maître assistant
 DR Directeur de recherche
 Ing. Ingénieur
 MR(DR2) Maître de recherche
 CR Chargé de recherche
 EC Enseignant-chercheur
 IGM Ingénieur général des mines

Dernière mise à jour le : 9 mars 2010

Centres :

SMS Sciences des Matériaux et des Structures
 SPIN Sciences des Processus Industriels et Naturels
 SITE Sciences Information et Technologies pour l'Environnement
 G2I Génie Industriel et Informatique
 CMP Centre de Microélectronique de Provence
 CIS Centre Ingénierie et Santé

« Il y a un but, mais pas de chemin; ce que nous nommons chemin est une hésitation ».
Franz Kafka.

Avant-propos

Cette thèse est le fruit de la collaboration entre le département Packaging et Supports Souples du Centre de Microélectronique de Provence – Georges Charpak de l'Ecole Nationale Supérieure des Mines de Saint-Etienne et l'équipe de Simulation Mécanique et Thermomécanique (MTM) de ST Microelectronics à Crolles. Pour m'avoir accueilli au sein de leurs équipes respectives, je remercie Clément Tavernier, responsable du groupe TCAD et Patrick Benaben, responsable du département PS2.

Mes remerciements vont également à François Bay et Philippe Goudeau qui ont accepté la lourde tâche d'être les rapporteurs de ce manuscrit. De même, je tiens à remercier Sandrine Bec et Lucia Zullino qui ont accepté de faire partie du jury, et m'ont fourni des commentaires avisés concernant les différentes parties de la thèse.

Au sein de ST Microelectronics, je tiens à remercier les personnes qui, à Crolles ou à Grenoble, m'ont aidé à réaliser les différents travaux inclus dans cette thèse. Tout d'abord, merci à Jean-Pierre Gonchond pour m'avoir permis de mieux appréhender la nanoindentation. Merci également à Géraldine Thiery, Véronique Guyader, Frédéric Lorut, Philippe Delpech et Fabien Quercia pour m'avoir soutenu dans la mise en œuvre et la réalisation des différentes parties expérimentales concernant le « probing » et le « wire bonding ».

Pour les moments passés à Gardanne, je remercie ceux qui m'ont aidé à avoir une vision différente de certains aspects de ma thèse.

Je remercie également mes « partenaires » de repas et pauses que ce soit à Crolles (Mehdi, Alban, Sébastien, Olivier, Pierre, François, Lionel...) ou à Gardanne (Renaud, Thomas, Etienne, Mohamed, Romain(s), Brice, Jessica...) pour les discussions scientifiques mais également les moments plus détendus.

Un thésard peut difficilement effectuer correctement son travail s'il n'est pas bien encadré. Ainsi, au sein de ST Microelectronics, je tiens à remercier tout d'abord Florian Cacho, qui a eu la patience de m'initier à la simulation, et Vincent Fiori qui m'a ensuite suivi et soutenu afin de mener à bien ces travaux. Leurs disponibilités et leurs compétences, tant dans le domaine de la microélectronique que dans celui de la modélisation numérique, m'ont été d'une aide précieuse afin de ne pas perdre de vue les objectifs de la thèse. Je tiens également à remercier Karim Inal pour son soutien, les nombreuses discussions scientifiques en particulier en ce qui concerne la mécanique des matériaux. La distance ne s'est finalement que peu faite sentir au cours de ces années. Je remercie aussi Xavier Boddaert qui, malgré le fait que la thèse soit éloignée de ses domaines de prédilection, m'a apporté un soutien ponctuel et efficace.

Je remercie enfin tous les personnes de mon entourage, mes parents, ma sœur pour leur soutien chaleureux et indéfectible ainsi que mes amis de Grenoble (Guillaume, Thibaut, Micky, Antoine...) ou des divers coins de l'Europe (Alexis, les ingénieurs « VC&T »,...) pour les très agréables moments de détente qui ont jalonné cette thèse.

Comme toute série de remerciements, il y a probablement des personnes omises, que celles-ci veuillent bien m'excuser.

General introduction

I.	Industrial context.....	- 9 -
II.	Solving approach.....	- 10 -

Chapter 1: Downscaling from IC to pad: description, elaboration, related processes and mechanical issues

I.	Introduction: integrated circuits and pads	- 12 -
I.A.	Integrated circuits: description	- 12 -
I.A.1.	Generalities on IC.....	- 12 -
I.A.2.	Front end	- 12 -
I.A.3.	Back end.....	- 14 -
I.B.	Description and elaboration of the BEoL and the pad.....	- 15 -
I.B.1.	Description of the interconnections.....	- 15 -
I.B.2.	Elaboration of the oxide and copper layers	- 17 -
I.C.	Introduction to mechanical issues in Pads	- 18 -
II.	Electrical wafer sort:	- 19 -
II.A.	Test description.....	- 19 -
II.A.1.	Introduction to EWS process	- 19 -
II.A.2.	Test description and parameters	- 20 -
II.A.3.	Probe Cards.....	- 20 -
II.A.4.	Electrical measurements and cleaning.....	- 23 -
II.B.	Failure analysis: pad integrity	- 24 -
II.C.	Introduction to mechanical issues during probing	- 25 -
III.	Wire bonding and related testing	- 26 -
III.A.	Test description.....	- 26 -
III.A.1.	First techniques to bond wires	- 26 -
III.A.2.	Thermosonic ball bonding	- 27 -
III.A.3.	Materials in wire bonding	- 28 -
III.B.	Bond characterization tests	- 32 -
III.B.1.	Shear test.....	- 32 -
III.B.2.	Wire bond pull test	- 34 -
III.B.3.	Thermal aging.....	- 36 -
III.C.	Mechanical issues during wire bonding.....	- 36 -
IV.	Other processes inducing mechanical failures in pads.....	- 36 -
IV.A.	Flip Chip assembly	- 37 -
IV.A.1.	Solder bump.....	- 37 -
IV.A.2.	Copper pillar	- 37 -
IV.B.	Toward 3D integration: TSV	- 38 -
V.	Problematic and conclusions	- 38 -
References	- 40 -

Chapter 2: Comprehension of the mechanical failures: Presentation of the experimental and numerical tools set up

I.	Experimental techniques for failure analysis	- 46 -
I.A.	3D observation of the pad surface: Optical profilometry	- 46 -
I.A.1.	Principle	- 46 -

I.A.2.	Expectations and limitations of the technique.....	47 -
I.B.	3D reconstruction of the pad volume: FIB/SEM reconstruction	47 -
I.B.1.	FIB and SEM principles	48 -
I.B.2.	Advantages and drawbacks of the technique	48 -
I.C.	Additional tools used to evaluate the pad damages	49 -
I.C.1.	Electron backscatter diffraction	49 -
I.C.2.	Energy dispersive X-Ray	49 -
II.	Experimental evaluation of the mechanical robustness of pads: nanoindentation	50 -
II.A.	Principle and apparatus	50 -
II.A.1.	Presentation of the technique.....	50 -
II.A.2.	Berkovich and cube corner indenter tips	51 -
II.A.3.	General description of the indentation process.....	52 -
II.B.	Determination of material properties	53 -
II.B.1.	Measurement of hardness and elastic modulus	53 -
II.B.2.	Oliver and Pharr model and limitations.....	53 -
II.B.3.	Parameters affecting the indentation curves	55 -
II.B.4.	Continuous stiffness measurement	56 -
II.C.	Considerations for the indentation of multilayer structures.....	56 -
II.D.	Nanoindentation experiments for pad robustness evaluation	58 -
III.	EWS numerical model	58 -
III.A.	Model description	59 -
III.A.1.	Geometrical model	59 -
III.A.2.	Material Properties	60 -
III.A.3.	Loading conditions and contact checking	61 -
III.A.4.	Physical and numerical parameters	62 -
III.B.	EWS numerical model set up: calibration	66 -
III.B.1.	Experiment and calibration for CMOS 120nm pad model.....	67 -
III.B.2.	Experiment and calibration for CMOS 45nm pad model.....	71 -
III.B.3.	Modal analysis of the probe behaviour.....	74 -
IV.	Conclusions	77 -
	References	79 -

Chapter 3: Test and assembly critical parameters on the appearance of mechanical issues in pads

I.	EWS tests parameters effect on the crack occurrence.....	82 -
I.A.	Dynamical effects during probing	82 -
I.A.1.	Chuck velocity.....	82 -
I.A.2.	Chuck vibrations	84 -
I.B.	Test parameters effect: overdrive and number of passes.....	84 -
I.B.1.	Observation and measurements of the probe mark	85 -
I.B.2.	Pad integrity results and observations	88 -
I.B.3.	Finite element model results and limitations.....	89 -
I.B.4.	Discussion	90 -
II.	Gold bonding: Influence of the process parameters.....	91 -
II.A.	Gold bonding with 25µm diameter wires	91 -
II.A.1.	First bonding Design of Experiments	92 -
II.A.2.	Second wire bonding DOE	95 -
II.A.3.	Discussion.....	96 -
II.B.	Gold bonding with 20µm diameter wires	97 -

II.B.1.	Bonding parameters effects on mashed ball dimension	- 97 -
II.B.2.	Wire pull tests results	- 98 -
II.B.3.	Ball adhesion on pads	- 98 -
II.B.4.	Discussion.....	- 99 -
II.C.	Effect of the thermal aging	- 99 -
II.C.1.	Wire pull tests results	- 100 -
II.C.2.	Ball adhesion on pads	- 100 -
II.D.	Discussion	- 101 -
III.	Copper bonding: effect of the process parameters	- 101 -
III.A.	Initial bonding observations and copper ball measurements	- 102 -
III.A.1.	Aluminium pad splash	- 102 -
III.A.2.	Mashed ball measurements.....	- 103 -
III.B.	Wire pull tests results.....	- 103 -
III.C.	Shear tests results.....	- 105 -
III.C.1.	Effect of the shear test parameters on forces	- 105 -
III.C.2.	Effect of the contact force and the bonding power.....	- 106 -
III.D.	Discussion	- 108 -
III.D.1.	Comparison between gold and copper bonding	- 108 -
III.D.2.	Weight of the process and design parameters	- 109 -
IV.	Conclusions	- 110 -
	References	- 111 -

Chapter 4: Layout effects on the mechanical robustness of pads

I.	Stacking influence: thickness of the top layers	- 116 -
I.A.	Failure analyses after EWS.....	- 117 -
I.A.1.	Crack occurrences after pad integrity tests	- 117 -
I.A.2.	3D reconstruction of the cracks under the pads	- 118 -
I.B.	Resistance to compression: nanoindentation tests.....	- 120 -
I.B.1.	Cube corner tip	- 120 -
I.B.2.	Berkovich tip	- 125 -
I.C.	EWS numerical model: stress analysis.....	- 129 -
I.D.	Discussion and correlation between techniques about the thickness effect	- 130 -
I.D.1.	Comparison between techniques to evaluate the thickness effect.....	- 130 -
I.D.2.	Discussion	- 131 -
II.	Effect of the pad layout: metal density, presence of vias	- 131 -
II.A.	Failure observed after wire bonding tests	- 132 -
II.A.1.	Pad description	- 132 -
II.A.2.	Pull tests.....	- 133 -
II.A.3.	Shear tests	- 136 -
II.A.4.	Discussion.....	- 137 -
II.B.	Nanoindentation tests on CMOS 120nm	- 139 -
II.B.1.	Nanoindentation with Berkovich tip.....	- 139 -
II.B.2.	Nanoindentation with cube corner tip	- 139 -
II.B.3.	Discussion: wire bonding versus nanoindentation	- 144 -
II.C.	Nanoindentation on CMOS 45 nm pads	- 145 -
II.C.1.	Pad description	- 145 -
II.C.2.	Measurements with Berkovich tip.....	- 146 -
II.D.	Discussion on the pad layouts effect.....	- 148 -
III.	Metal copper layer design	- 148 -

III.A.	Pad layouts description	148 -
III.B.	Pad integrity tests	149 -
III.C.	Nanoindentation	151 -
III.C.1.	Berkovich tip	151 -
III.C.2.	Cube corner tip	151 -
III.D.	Numerical results	152 -
III.E.	Discussion	153 -
IV.	Nanoindentation and modelling to reproduce EWS and wire bonding.....	155 -
V.	Conclusions	156 -
	References	158 -

Chapter 5: Enhancement of the experimental and numerical methods and tools to evaluate the robustness of pads

I.	Nanoindentation: pad robustness qualification	164 -
I.A.	Comparison between Berkovich and cube corner tips	164 -
I.A.1.	Dimension and shape of the indent	164 -
I.A.2.	Influence of the tip geometry on the load-displacement curves.....	165 -
I.A.3.	Determination of the Young's modulus and hardness	166 -
I.B.	Application to the determination of the pad robustness	168 -
II.	Modelling: FEM dedicated to pad enhancement to reproduce EWS	169 -
II.A.	Comparison between explicit and implicit methods	169 -
II.A.1.	Time integration	169 -
II.A.2.	Stability limits.....	170 -
II.B.	Development of the 2D implicit model	170 -
II.B.1.	Objectives and use of experimental and numerical insights.....	170 -
II.B.2.	Model description	171 -
II.B.3.	Loading conditions of the 2D model	172 -
II.B.4.	Results and limitations of the 2D FE model.....	176 -
II.C.	Development of the 3D implicit FE model.....	179 -
II.C.1.	Description of the 3D FE model of the pad.....	179 -
II.C.2.	Loading conditions	179 -
II.C.3.	Results of the 3D model	180 -
II.C.4.	Insertion of an elastic-plastic behaviour for the aluminium	182 -
II.D.	EWS simulations: explicit versus implicit 3D modelling.....	187 -
III.	Conclusions	187 -
	References	189 -

Conclusion and perspectives

I.	Synthesis.....	191 -
II.	Perspectives	193 -

Appendixes

A.	Comparison between nanoindentation, EWS and wire bonding.....	196 -
A.1.	Pad layout description	196 -
A.2.	EWS results	197 -
A.3.	Nanoindentation tests	197 -
A.3.1.	Results with the Berkovich tip	197 -

A.3.2. Results with the cube corner tip	197 -
A.4. Wire bonding	198 -
A.5. Correlations between techniques	198 -
B. DOE used for the evaluation of the test and assembly parameters and the pad layout designs	199 -
C. Acronyms	200 -

Résumé

I. Description des tests et processus induisant des défaillances dans les pads	204 -
I.A. Tri électrique des plaques (EWS)	204 -
I.A.1. Description du test	204 -
I.A.2. Défaillances associées et méthodes d'observation	205 -
I.B. Assemblage filaire	206 -
I.B.1. Description du processus	206 -
I.B.2. Méthodes de caractérisation	206 -
I.C. Problèmes mécaniques induits	207 -
II. Présentation des méthodes expérimentales et numériques mises en œuvre	208 -
II.A. Observation et reconstruction du pad en 3 dimensions	208 -
II.B. Evaluation de la robustesse des pads : nanoindentation	208 -
II.C. Modèle numérique reproduisant l' EWS	209 -
III. Effet des paramètres de test et d'assemblage sur l'apparition des défaillances	211 -
III.A. Paramètres critiques lié au test EWS	211 -
III.B. Connexion électrique par fil de cuivre	212 -
IV. Effet de l'architecture des pads sur robustesse mécanique	214 -
IV.A. Effet de l'épaisseur des couches supérieures	214 -
IV.B. Effets de la densité de cuivre et de la présence des vias	218 -
IV.C. Effet de l'agencement de la couche de cuivre	221 -
V. Adaptation au contexte industriel	223 -
V.A. Méthodologie pour l'étude des architectures de pad	223 -
V.B. Développement d'un nouveau modèle numérique reproduisant l'EWS	225 -
VI. Conclusions	227 -
Références bibliographiques	229 -

General Introduction

I. Industrial context

In the semiconductor industry, and more specifically at STMicroelectronics, the forward move to reduce the critical dimensions in the semiconductor devices is accompanied by the introduction of new integration ways (3D integration) and the use of new brittle materials, such as low-k oxides. These materials are required to enhance the insulating capabilities of the dielectric layers, but they weaken the mechanical structure of the pads.

During the manufacturing of an integrated circuit, parametric tests such as the Electrical Wafer Sort (EWS) and assemblies like the wire bonding are performed on the pads at the end of the front end process. The EWS, also known as probing, is required to check the electrical functionality of the die at the wafer level. It consists in carrying out a contact between a stiff probe and the aluminium bond pad. The assemblies are required to connect the die to the outer part before the packaging steps. The wire bonding, which is currently the most common way to achieve the electrical connection, is facing an evolution due to the increasing cost of gold. Thus, there is an interest in moving toward copper wires to achieve the bonding. However, the use of copper raises several questions due to the increase of the bonding process parameters necessary to weld the wire to the pad.

As a consequence, during these manufacturing steps, the pad and the interconnections below undergo heterogeneous stresses, which can lead to failures (cracks in the oxide during probing and peeling during the wire pull tests following the bonding). These failures induce electrical issues such as shorts circuits. Hence, the integrity of the bond pad becomes a major concern for technology integration and yield loss.

Thus, the main objectives of this work are to control the tests and process parameters in relation with the optimization of the pad architectures. To that purpose, some experimental and numerical techniques are used to gain insights on the failures in the interconnections, their location and root causes, and then reproduce the actual EWS and wire bonding. The driving force to mimic EWS tests and wire bonding is the duration and expensiveness of large designs of experiments, both for test and process parameters and for pad architecture optimization. As a matter of fact, these studies require patterned silicon wafers, the availability of the tools (probers, bonders...) and also dedicated equipment (probe card dedicated to the wafer for the EWS) and materials (copper or gold wires for wire bonding). Thus, in an industrial context, there is a need to have faster and cheaper reliable tools for technology development or crisis resolution. In that frame, nanoindentation and finite element modelling are particularly adequate to try to reproduce the process and test loading conditions. The former is a simplified loading condition (compressive test) with respect to the EWS and wire bonding, which only requires dies. The latter only requires the development and calibration of a numerical model, and allow the investigation of a large amount of test conditions and structures.

The complexities of the evaluation of the EWS and the bonding operations on pads are multiple. Indeed, the study of test and process parameters requires the evaluation of the inertial effects during these operations, and more specifically the contact between the needle tip and the bond pad. Then, this is a multi-scale issue from the size of the interconnections (a few micrometres) to the size of the operating tools (e.g. the needle for probing is in the range of few millimetres). Moreover, the pad is a multilayer system with complicated design and very different material behaviours that undergoes high damages during these operations and sometimes failures. Hence, the evaluation of the pads and the reproduction of the test and process must take into account all these considerations to be relevant.

II. Solving approach

After introducing the manufacturing steps of an integrated circuit, the electrical wafer sort, the assemblies and their relative mechanical issues are briefly presented. In the second chapter, the characterization techniques to gain insight on the failures occurring in the interconnections below the pad are detailed. Then, the nanoindentation, which is a recent technique used to characterize thin films, and the finite element model (using explicit scheme) set up to mimic the EWS test and assembly are described. The advantages and expected limitations in the reproduction of the loading conditions of the actual tests are also questioned.

In the next chapters, several designs of experiments, based on many test, process and pad design parameters are performed to improve the understanding of the failure mechanisms. Thus, in the third chapter, the objective is to determine which test and process parameters have an influence on the damaging of the top aluminium layer and the appearance of failures in the oxide layers of the interconnections. After that, some rules to control the test and process parameters are expected in order to reduce the fracture hazard.

The fourth chapter is dedicated to the evaluation of the influent pad design parameters on the mechanical robustness of the structure. Indeed, several layouts at various scales of the back end of line level (layer thicknesses, metal density, presence of via and layout of a specific metal level) and also some interactions with the failures (cracks during EWS, peeling after wire pull tests) are investigated. The determination of the critical loading conditions prior to the appearance of the failures for the various designs is also investigated. Then, correlations between the experimental and numerical techniques using adequate indicators are questioned. Finally, some guidelines for the optimization of the pad mechanical robustness are expected.

As a work in an industrial context, there is a need to have fast and reliable nanoindentation and finite element model to reproduce the loading conditions of EWS and wire bonding and to evaluate various pad structures. In the last chapter, some rules for the comparison of the mechanical robustness of pads using the nanoindentation technique are clarified. Afterwards, an implicit finite element model based on all the insights of the experiments and simulations previously achieved and dedicated to the pad study is set up.

Chapter 1:

Downscaling from IC to pad: description, elaboration, related processes and mechanical issues

Outline

I.	Introduction: integrated circuits and pads	12 -
I.A.	Integrated circuits: description	12 -
I.A.1.	Generalities on IC.....	12 -
I.A.2.	Front end	12 -
I.A.3.	Back end.....	14 -
I.B.	Description and elaboration of the BEoL and the pad.....	15 -
I.B.1.	Description of the interconnections.....	15 -
I.B.2.	Elaboration of the oxide and copper layers	17 -
I.C.	Introduction to mechanical issues in Pads	18 -
II.	Electrical wafer sort:	19 -
II.A.	Test description.....	19 -
II.A.1.	Introduction to the EWS process	19 -
II.A.2.	Test description and parameters	20 -
II.A.3.	Probe Cards.....	20 -
II.A.4.	Electrical measurements and cleaning.....	23 -
II.B.	Failure analysis: pad integrity	24 -
II.C.	Introduction to mechanical issues during probing	25 -
III.	Wire bonding and related testing	26 -
III.A.	Test description.....	26 -
III.A.1.	First techniques to bond wires	26 -
III.A.2.	Thermosonic ball bonding	27 -
III.A.3.	Materials in wire bonding.....	28 -
III.B.	Bond characterization tests	32 -
III.B.1.	Shear test.....	32 -
III.B.2.	Wire bond pull test	34 -
III.B.3.	Thermal aging	36 -
III.C.	Mechanical issues during wire bonding.....	36 -
IV.	Other processes inducing mechanical failures in pads.....	36 -
IV.A.	Flip Chip assembly	37 -
IV.A.1.	Solder bump.....	37 -
IV.A.2.	Copper pillar	37 -
IV.B.	Toward 3D integration: TSV	38 -
V.	Problematic and conclusions	38 -
References	40 -

I. Introduction: integrated circuits and pads

In this first chapter, the various scales set up in the semiconductor industry are swept: from the transistor to the size of a moulded chip. A global description of the several hundred steps required to build the chip is performed. Then, a focus is made on the interconnections and the pads, which allow the electrical connections with the outer part of the integrated circuit. The probing, wire bonding and other assemblies during the chip processing are realized on pads during the packaging operations. These tests and assemblies and their induced mechanical issues, failures in the oxide layers, are widely described.

I.A. Integrated circuits: description

I.A.1. Generalities on IC

The Integrated Circuits (IC) are manufactured on silicon wafers. Nowadays, these silicon wafers have usually 200mm or 300mm diameter and are made from silicon single crystal. Silicon, as a semi conductor material, has the electrical properties of conductors and insulators materials depending on its doping. Hence, electrical conduction is possible under specific conditions, and can be tuned.

The manufacturing of I.C. is usually divided in 2 parts:

- The Front End (FE) composed of electrical compounds and interconnections.
- The Back End (BE), which refers to the sawing, the assembly and the packaging.

I.A.2. Front end

The front end is composed of the electrical compounds and the interconnections as depicted in the Figure I.1.

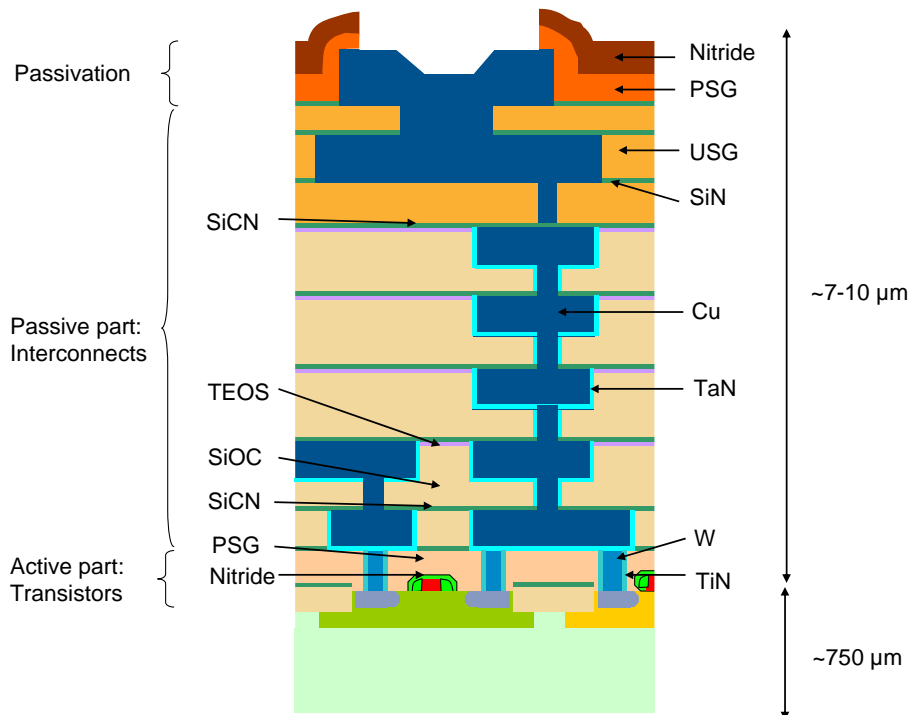


Figure I.1. Side view of the FE part of a chip for a CMOS 90nm technology node.

The electrical compounds represent the active part of the FE, also known as the Front End of Line (FEoL). Usually, these are transistors, but resistances, capacitors and diodes are found as well. The interconnections are the passive part of the FE. The main purpose of this part of the I.C. is to connect electrically the transistors between themselves. The whole FE processing requires hundreds of operations.

1.A.2.a) FEoL: Transistors

There are two kinds of transistors, the bipolar and Metal Oxide Semiconductor (MOS). Most of the integrated circuits are based on MOS transistor type and its applications are wide such as counting, memorizing and realizing logical functions.

MOS transistors are always 4 terminal devices, which include a gate, a source and a drain (doped semiconductor areas) and the bulk (silicon). The source and the drain are inversely doped with respect to the bulk. Therefore, two types of MOS transistors are distinguished: the NMOS (source/drain doped N with atoms such as P, As...) and the PMOS (doping with B, In...). The gate is separated from the substrate by an oxide layer. The functioning principle of a NMOS transistor is the following: by applying an electric field, the channel resistance can be modulated. More precisely, under the application of a positive tension to the gate, the electrons of the N-doped areas (source and drain) are attracted under the gate to form an electron channel, as shown in the Figure I.2. On the contrary, the holes from the silicon bulk are repulsed. This channel is obtained when a threshold tension is reached, allowing the electrical current to pass from the source to the drain while a tension is applied between both parts. The functioning of the PMOS is the same but with inversed batteries, and the channel is made from holes. Several NMOS and PMOS transistors are usually linked to realize logical functions, and called Complementary MOS (CMOS).

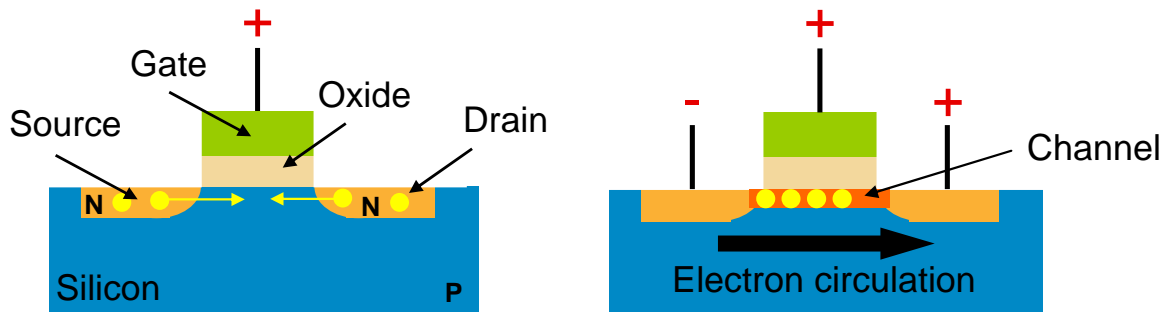


Figure I.2. NMOS type transistor functioning: positive tension is applied to the gate inducing the attraction under the gate of the electrons (left) to form a channel, while a threshold is reached, that can further allow the electron circulation from the source to the drain (right).

The minimum width of the gate determines the technology node. This width regularly decreases following the well-known Moore's law. Nowadays, the industry is able to produce technologies down to 45nm (CMOS 45nm or C45), nevertheless 65nm (CMOS 65nm), 90nm (CMOS 90nm) and 120nm (CMOS 120nm) gate widths are still commonly industrially produced for specific applications.

1.A.2.b) BEoL: Interconnections

The interconnections correspond to the passive part of the integrated circuit, also called Back End of Line (BEoL). Their purpose is to link the electrical compounds to create the desired functionality for the chip. These are made from several materials (see Figure I.1), whose major ones are metals and oxides. Metal layers and vias are currently made from copper and ensure the electrical connections. To electrically insulate these layers, some

dielectric materials are used. Several kinds of materials can be used for insulation. They are later detailed in the sections I.B and I.C. At the end of the FE process, before performing BE steps, a parametric test is performed to check the functionality of the chip: the electrical wafer sort.

I.A.3. Back end

The back end part is composed of all the steps from the thinning of the wafer to the end of the packaging. As described by the Figure I.3, the die manufacturing consists in four main successive steps: the thinning, the sawing, the assembly and finally the packaging. The Figure I.4 presents an integrated circuit at the end of the back end steps, opened on one quarter, to show the chip and the assembly.

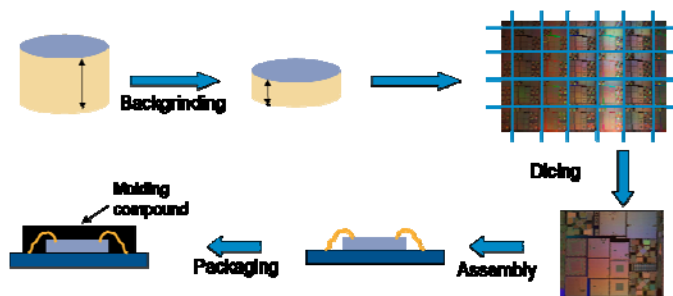


Figure I.3. Description of the back end main steps.

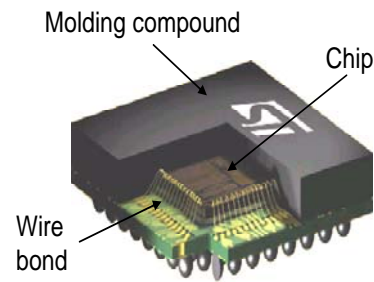


Figure I.4. View of an IC with partial opening of the moulding.

Various back end processes exist to ensure the electrical connection between the chip and the packaging of the die. In the next section, only the major ones are described. Many mechanical issues (delamination, die fracture...) are raised during all these back-end stages. In this work, attention is only paid to those occurring in the interconnections.

I.A.3.a) Silicon backgrinding

This step is required to simplify the cut of the wafer and to reduce the final size of the chip. Wafer thinning is achieved by polishing the back side of the wafer from about $750\mu\text{m}$ down to $50\mu\text{m}$ depending on the die application.

I.A.3.b) Dicing

The purpose of this step is to separate the dies from a wafer. This process is accomplished by scribing, which is made either by mechanical sawing (dicing saw) or by laser cutting, and then breaking the wafer. In order to keep the electrical function undamaged, the dicing is performed in scribe lines.

I.A.3.c) Assembly

After being split, the dies are packaged. The chip is flipped (Flip Chip or FC) or not depending on the application, the number of dies packaged together and the required electrical connections. In each case, several assembly techniques exist, and only the most common ones are further described. Most of these techniques are currently applied and induce stress in the pad structure leading sometimes to failures. As a consequence, the assembly techniques and related issues are more widely treated in sections III and IV of this chapter.

For a standard assembly, within top side of the dies faces up, commonly used processes to link electrically the integrated circuit to the Printed Circuit Board (PCB) are the wire bonding and the use of Through Silicon Vias (TSV). Flip chip, within top side of the dies faces down,

also allows establishing a connection between the PCB and the die. For flip chip, the connections are made with solder bumps or copper pillars.

I.A.3.d) Encapsulation

Once the assembly performed, the die is encapsulated in order to protect the chip from physical damages or corrosion effects. This step is achieved by using moulding compounds, which can be ceramics, plastics or epoxy resins. An epoxy resin matrix is often used as filler for the encapsulation.

I.B. Description and elaboration of the BEOl and the pad

I.B.1. Description of the interconnections

The pad refers to the area of the integrated circuit not covered by the passivation layers (nitride and oxide layers). The opening is performed at the end of the FE steps. As a consequence, the pad is located on top of the interconnections and is composed of the BEOl and the aluminium capping on top of it. In this volume, the electric connections between the transistors and with the outer part of the I.C. are set up. Thus, good electrical properties and mechanical resistance are required, since the parametric test (EWS) and assemblies are made on top of the pads.

Nowadays, the metal layers from the BEOl are made from copper. Their number varies from 4 to 7, depending on the technology node as depicted in Figure I.5.

For latest technology nodes, the top of the pad is made from an aluminium alloy (99.5% of aluminium and 0.5% of copper) to enhance its mechanical robustness. Moreover, this aluminium alloy (also often referred as alucap) is related to lower corrosion issues compared to copper. Indeed, a 2 to 5 nm native aluminium oxide is formed on its surface when exposed in air. This very thin layer is thermodynamically stable and its thickness varies slowly with time and heat. The aluminium oxide has an amorphous structure that protects aluminium for further oxidation. However, this oxide barrier is an excellent electric insulator (support electric field in the order of 1V/nm).

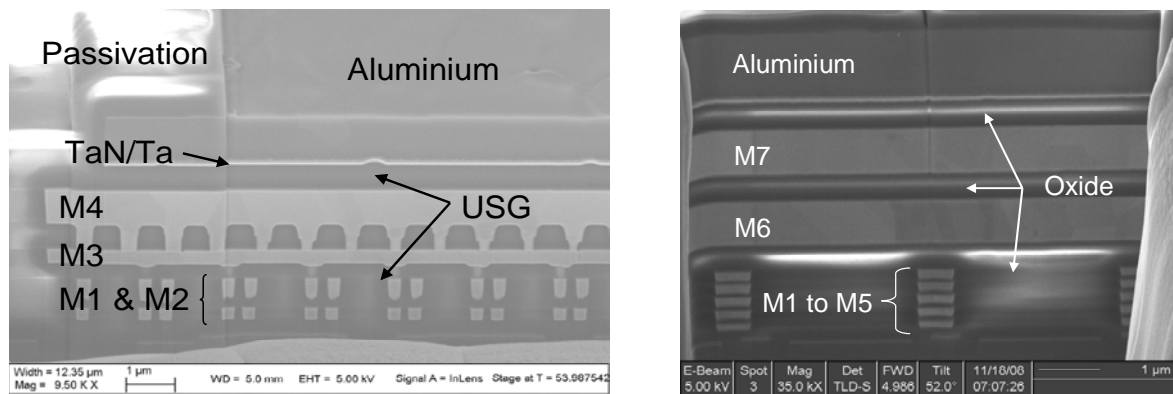


Figure I.5. Tilted SEM side view of a pad for CMOS 120nm (left) and for CMOS 45nm (right).

The other major family of materials found in the BEOl is the oxides, also known as Inter-Metal Dielectric (IMD). This family has insulation purposes and evolves with technology nodes. This evolution is necessary since thinner layers are used in the bottom part of the BEOl, requiring higher insulation capabilities from the oxides to lower the capacitances. These insulating capabilities are defined by the dielectric constant (noted k). The dielectric constant of SiO₂ is 4. Low-k materials have dielectric constants below 4 that can reach values close to the air (see Table I.1.). Thus, 3 main sub-families can be distinguished, with microelectronic applications:

- Silica based (see Figure I.6): these oxides have the tetrahedral basic structure of SiO_2 . They are either undoped such as USG (Undoped Silica Glass) or doped like FSG (Fluorinated Silica Glass). The latter are obtained from TEOS ($\text{Si}(\text{OEt})_4$), with k values between 3.2 and 4. Another approach is the introduction of carbon, through CH_3 groups, in silicate glasses. The k value of the material depends on the number of CH_3 groups and is in the range 2.6-3. This family is by far the most common in current technology nodes and allows obtaining low- k materials by introducing porosities (see Figure I.8).
- Silsesquioxane based (SSQ): the elementary unit is silsesquioxane $(\text{R-SiO}_{3/2})_n$, as depicted in Figure I.7. These are organic-inorganic polymers and allow obtaining lower dielectric constants. Most common forms are hydrogen-silsesquioxane (HSQ) and methyl-silsesquioxane (MSQ, commercially known as Black Diamond®), whose dielectric constants are respectively 3-3.2 and 2.8.
- Organic polymers: these are polymers with C-C polar or non-polar bonds. Dielectric constants can reach 2.6-2.8, but many thermal stability issues can be encountered and some incompatibility with process limits their applicability.

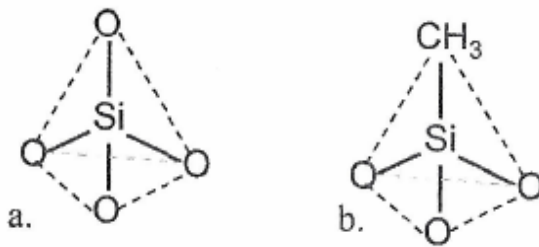


Figure I.6. Elementary units of: (a) SiO_2 and (b) C doped silica glass.

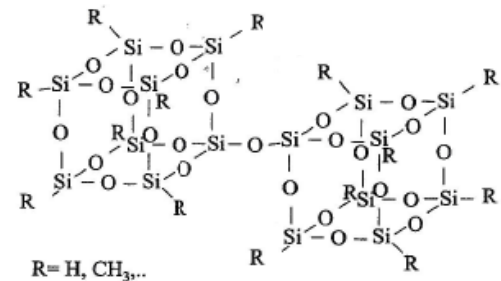


Figure I.7. Structure of elementary units of SSQ dielectric materials.

Dielectric materials	Dielectric constant
SiO_2	3.9-4.1
SiOF	3.2-4.0
Polysiloxane	2.7-4.0
Polyimide	2.9-3.9
Fluorinated polyimide	2.7-3.0
Fluoropolymers	1.8-2.5
Aerogels and microporous polymers	1.1-1.8

Table I.1. Examples of dielectric materials and their dielectric constants after [Leu, 1995].

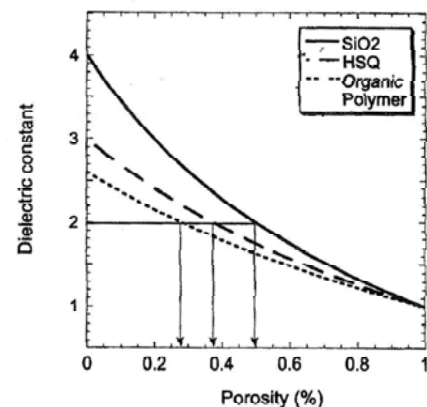


Figure I.8. Evolution of the dielectric constant with increasing porosity [Maex, 2003].

In latest technology nodes (CMOS 65nm and CMOS 45nm), the oxide deposited in the bottom of the BEoL belongs to the low- k family (X and Y levels, e.g. below the M5 layer for CMOS 45nm). The oxide of the top of the BEoL is USG, since thicker layers are deposited.

Apart from metal and oxide layers, other materials are used in the pad structure. These have specific functions during the manufacturing or for the chip reliability:

- TEOS is used in latest technology nodes to protect the low- k layers from Chemical Mechanical Polishing (CMP).
- SiCN, a silicon nitride layer doped with carbon, is used to avoid the copper oxidation during the process steps.

- TaN/Ta, which is very thin (30nm to 40nm), is deposited to avoid the copper diffusion in the oxide layers surrounding.

On top of the chip, 2 layers are deposited for protection purposes:

- PSG (Phosphorus Silica Glass) that acts as a diffusion barrier to metallic and organic contaminants.
- Silicon nitride (Si_3N_4) is used for mechanical protection and as a shield against humidity.

I.B.2. Elaboration of the oxide and copper layers

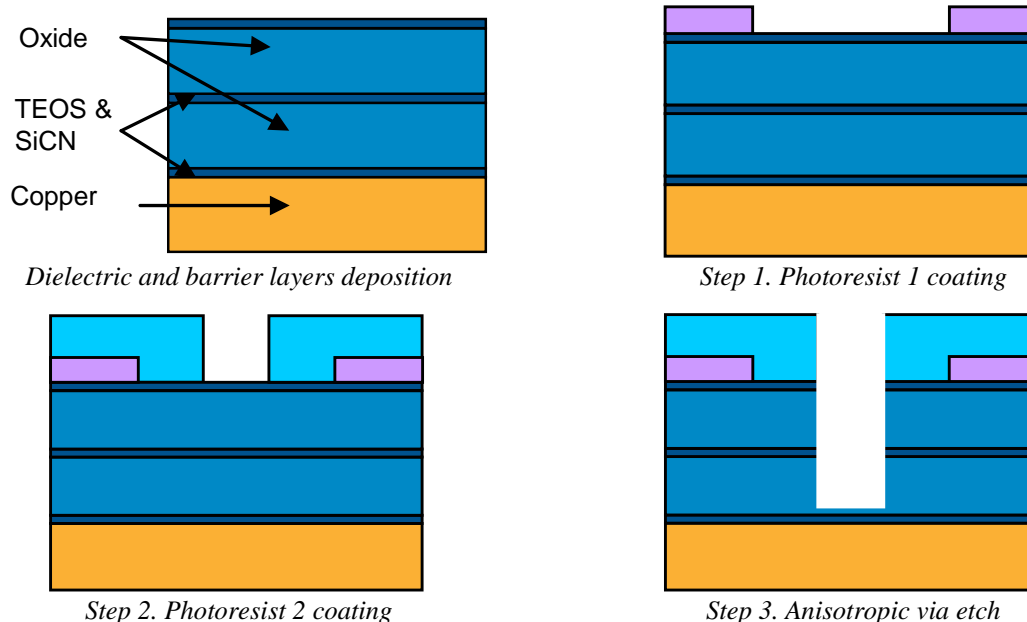
I.B.2.a) Dielectric layers

The oxide of the IMD can be deposited in 2 main ways:

- Chemical Vapour Deposition (CVD) is used for dielectric materials. A solid material is deposited on a substrate with the help of a gaseous precursor reacting on the substrate. Plasma could be used to reduce the deposition temperature (Plasma Enhanced CVD – PECVD).
- Spin-on is used for the deposition of low-k oxides. For this technique, either a sol (a stable suspension) or a gel (result from the interaction between a sol and a solvent) is poured on the substrate that is in rotation. The thickness of the deposited material depends on the deposition duration and the rotation velocity.

I.B.2.b) Copper lines and vias

The copper lines and vias are deposited following the dual damascene process. This process can be achieved by two etches: the trench etch and the via etch. Only the latter is described here since it corresponds to the current methodology to coat the wafers (Figure I.9).



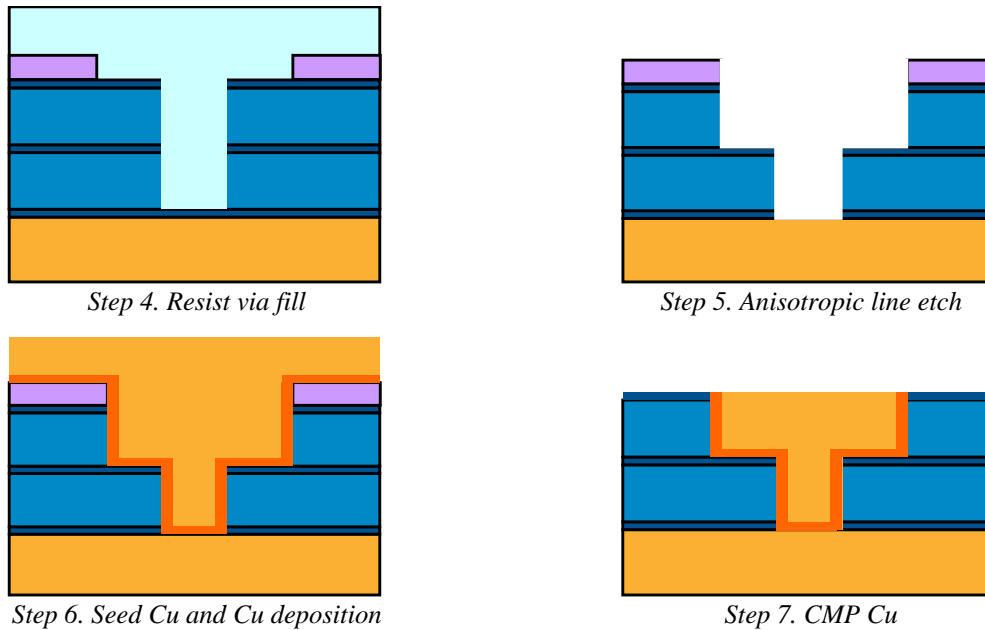


Figure I.9. Dual damascene process with via etch approach for the copper deposition.

Firstly, the wafers are coated with photoresist and lithographically patterned (steps 1 and 2). Then, anisotropic etching cuts through the surface hard mask and down to the IMD and the embedded etch stop to end on the bottom silicon nitride barrier (step 3). Next, the via mask is removed and the via is filled with resin (step 4). An anisotropic etching then cuts the surface hard mask and goes down through the oxide, stopping at the embedded hard mask. This etching forms the line (step 5). The photoresist is then stripped and the silicon nitride barrier at the bottom of the via is opened with a soft etching. TaN/Ta is deposited using Physical Vapour Deposition (PVD). To enhance the copper adhesion to the layer below, a Cu seed layer is deposited also by means of PVD. Copper can be deposited by means of electroplating (step 6). Finally, a thermal treatment and a Chemical Mechanical Polishing (CMP), which allows the surface planarization, are achieved (step 7).

I.C. Introduction to mechanical issues in Pads

The introduction of low-k oxides induces reliability issues. Indeed, the decrease of the dielectric constant is accompanied by a reduction of the mechanical properties (e.g. see Figure I.10: reduction of the elastic modulus). Characterisation studies to evaluate the mechanical properties of the dielectric thin films have been performed [Padovani, 2001; Vella, 2003; Volinsky, 2003; Brillet, 2007]. For most of the low-k oxides, the Young's modulus is below 12GPa and the hardness below 2GPa. Hence, these new materials have weak mechanical properties compared to those of silicon ($E=65-75\text{GPa}$, $H=7-9.5\text{GPa}$).

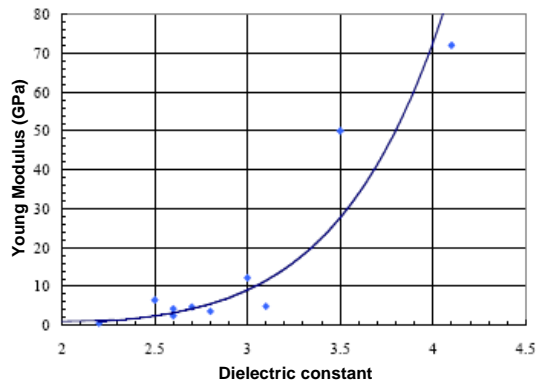


Figure I.10. Evolution of the Young's modulus as a function of the dielectric constant [Moore, 2001].

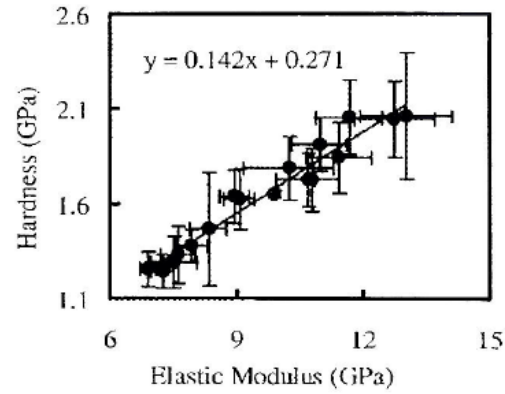


Figure I.11. Elastic modulus and hardness for organo-silicate glass (SiOCH family) [Vella, 2003].

Moreover, the adhesion energy values diminishes as well [Maex, 2003; Cherault, 2006]. Considering only the linear elastic domain, the crack propagation is directly related to the energy release rate, G . Hence, the lower the adhesion energy values, the easier the crack propagates. The adhesion energy value for low- k materials are between 2.5J/m^2 and 6J/m^2 , which is very low comparing to the ones of the silicon and its deriving materials (over 10J/m^2).

Mechanical issues are induced in silicon oxide during harsh or non adequate processes, tests or assemblies conditions. These failures are responsible for short circuits and lead to the reject of an integrated circuit and thus yield loss. In a more general manner, most of the FE/BE interactions implies mechanical stresses within the pad structure that can lead to failures in the interconnections. These parametric tests and assemblies, and the mechanical issues induced are described in the next chapters and sections (II.C, III.C and IV).

II. Electrical wafer sort:

Electrical wafer sort (EWS), also known as probing, is used to check the electrical functionality of the die at the wafer level. As mentioned previously, this test is performed at the end of the FE part of the process, before assembly.

II.A. Test description

II.A.1. Introduction to the EWS process

This test is part of the manufacturing process for over four decades. Patents on probing can be found as far as 1969 for microelectronic applications [Harmon, 1969]. To achieve the sort, a contact between the pad of a die and a contact element (needles, wires...) is established. This results in scrub marks (few tens of micrometres long) left in the aluminium layer on top of the pad.

At a larger scale, a prober, with electrical capabilities and software, is used to perform the electric tests and marks the rejected dies. It is composed of a test head, made from an epoxy ring containing the Probe Card (PC). The probe card contains several contact elements. On the other side, a chamber is used to store the wafers before and after being sorted. During the test, the prober holds the wafer in a vacuum chuck. The chuck moves in the horizontal direction to position the pads under the probe card and then, moves along the vertical axis to bring in contact the elements of the card and the pads.

II.A.2. Test description and parameters

During a test, a double touchdown is performed to improve the electrical measurements. It is required to remove the aluminium native oxide on top of the pad and to perform a reliable measurement. Many parameters should be taken into account while considering the test:

- The OverDrive (OD), which corresponds to the vertical displacement, in micrometres, applied to the chuck once the contact with the pad is established.
- The number of passes, which refers to the number of times the test is performed to get a good electrical response from the pad. Multiple passes are required to diminish the reject rate of dies.
- The velocity of the chuck during the test: on the one hand, a high test velocity fasten the overall testing time, on the other hand, it increases the odds to damage the pad, so a balance must be found.

One pass is made from several steps, as depicted by Figure I.12. At the beginning, the chuck is raised with a high velocity close to the probe card (about 10µm to the tip's end), and after a stabilization time, the first touchdown is performed at slower velocity. This first touchdown is composed of a loading step, followed almost directly by an unloading one. After another short stabilization time to avoid the oscillations of the needle and the chuck, the second touchdown starts. The loading ends when the OD value is reached, similarly to the first touchdown. Then, the electrical measurement is performed lasting a given time, and finally the unloading ends the pass.

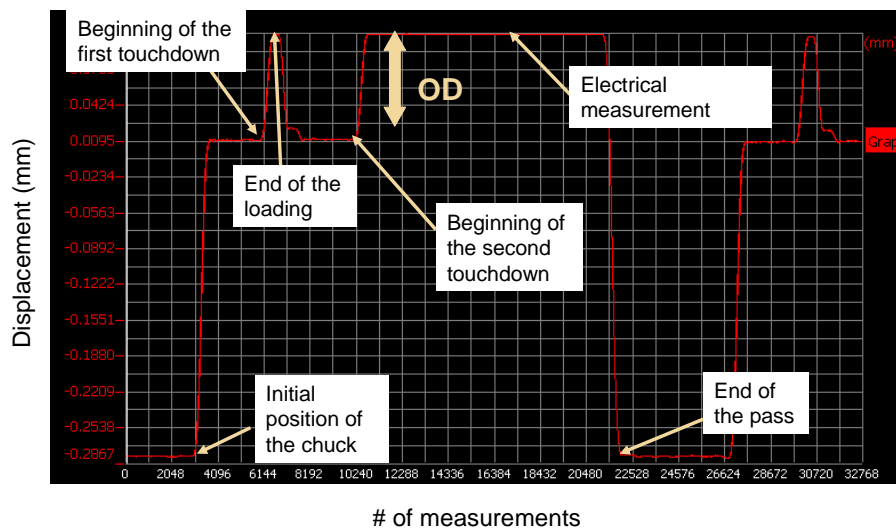


Figure I.12. Vertical movement of the chuck during a pass (2 touchdowns), measured by laser.

II.A.3. Probe Cards

The Probe Card (PC) holds the contact elements that establish the connection with pads and perform the electrical measurements. Three major kinds of PC exist depending on the I/O pad layout, and consequently to the die application: the cantilever probe card, the vertical probe card and the ones made from springs and/or MEMS (MicroElectronic Mechanical System). They are used respectively for conventional perimeter, area array and line on centre (see Figure I.13).

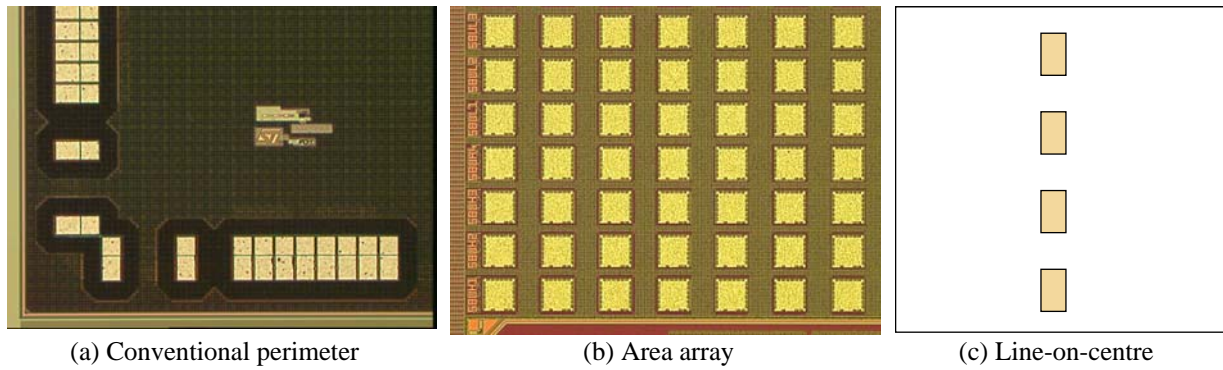


Figure I.13. Main I/O pad layouts. One pad has a size ranging from tens to hundreds of micrometres.

II.A.3.a) Cantilever probe cards

Nowadays, they are the most widely used probe cards. They are specifically designed to fit with conventional perimeter type of pad layout (Figure I.13 (a)). This type of probe card has been mainly developed in the 80's, with patents from Aigo [Aigo, 1984], Ohkubo [Ohkubo, 1986] or Evans [Evans, 1988]. One probe card has to be designed for each pad layout since needles should be accurately positioned on the pads.

As depicted by the Figure I.14, the probe card is composed of several needles (up to ten thousand). The end of these needles is attached to the probe card by means of an epoxy resin. Hard materials such as tungsten, tungsten-rhenium, beryllium-copper and other alloys are used to manufacture these needles. They are characterized by a bending angle, which is between 90° and 105° . Today, the minimum pitch between 2 needle tips is between $20\mu\text{m}$ and $30\mu\text{m}$ depending on the PC provider.

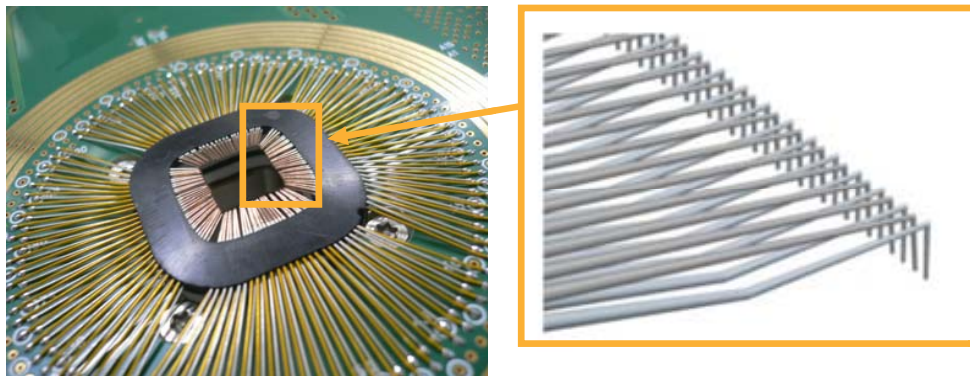


Figure I.14. Cantilever probe card (left) and set of needles (right) [Vallauri, 2007]

Cantilever probe cards are considered during further work and for most of the published papers on the probing topic. Hence, damages induced and additional information are discussed in chapters 3 and 4.

II.A.3.b) Vertical probe cards

The main alternative probe card is the vertical one. It best fits with area array type of pads (Figure I.13 (b)), although it can be used for conventional perimeter pad layout in some specific cases. The area array pad layout configuration is increasingly used since it is required for flip chip assemblies (see section IV.A for additional information on FC). In this design, the needles are attached to the probe card and are positioned almost vertically above the pads. This kind of PC has been firstly developed in the 70's by Bove [Bove, 1974] under the

buckling beam name because needles bend slightly during the test (see Figure I.15). Nevertheless, the most famous and promising one is the Cobra probe needle, patented by Byrnes and Wahl [Byrnes, 1977] as depicted by Figure I.16.

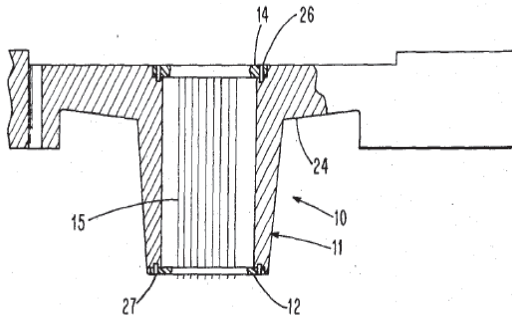


Figure I.15. Set of vertical needles [Bove, 1974].

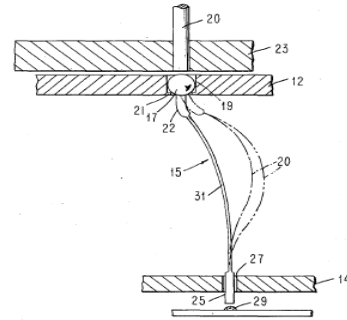


Figure I.16. Cobra needle [Byrnes, 1977].

The actual trend to develop the flip chip for some applications (high performance microprocessors and many high speed devices) increases the interest for these PC. As a consequence, some studies focuses on the design of the probe to reduce the stress during the bending of the needle [Chiu, 2009]. In another study [Hung, 2009], actual probing tests are performed on Cu/low-k structures (CMOS 65nm) and some design rules for the pad design are highlighted.

Vallauri et al. [Vallauri, 2007] performed a comparison between cantilever and vertical PC for a specific application (and for a conventional perimeter pad layout). Needles have buckling beam geometries and are mounted on specific cards (see Figure I.17). The comparison between the probe cards shows smaller scrub marks are expected while probing with the vertical one. Hence, the area of aluminium damaged by the electrical wafer sort can decrease by up to 50% as shown by Figure I.18 and Figure I.19. Moreover, according to this study, probing with buckled needles shows improved electrical contact and crack appears at higher loading.

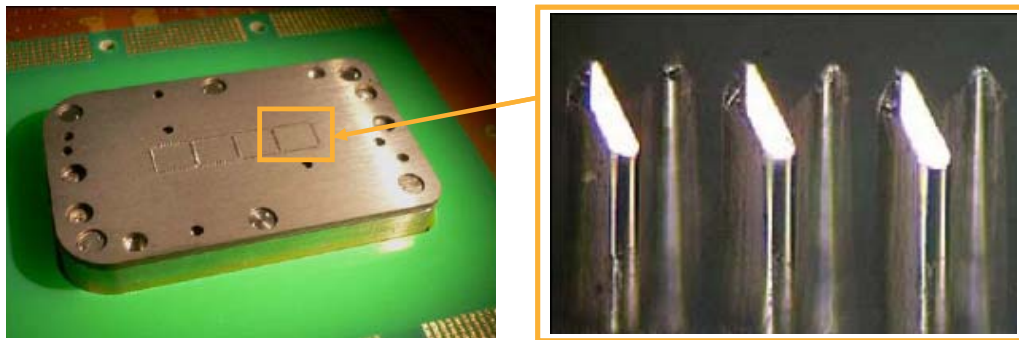


Figure I.17. Route60@ vertical probe card (left) and vertical needles tip set (right) [Vallauri, 2007].



Figure I.18. Scrub mark induced by probing with cantilever PC [Vallauri, 2007].



Figure I.19. Scrub mark induced by probing with Route60@ vertical PC [Vallauri, 2007].

Hauck et al. [Hauck, 2009] evaluated the forces and stresses induced by vertical probe card on pad structures in order to evaluate the significant design parameters, using finite element modeling. They found that buckling probe needles provide a constant probe force over a large range of overdrives. Pad designs like the number of vias in the top oxide layer and their distance to each other were also found to be critical parameters.

As previously described, the range of application for the vertical probe cards is widening. Nevertheless, the pitch for these PC is close to 60μm, which is higher than the pitch of cantilever PC, and can be limiting for some pad layouts.

II.A.3.c) Probe cards with micro-springs and/or MEMS

Some probe cards with micro-springs and/or MEMS are emerging. Figure I.20 and Figure I.21 depict the needles for such cards. Nowadays, they still have limited industrial applications but interest is paid to due to the reduction of the pitch size. Some recent studies evaluated various tip shapes and the damages induced while probing on pad structures with low-k oxides. Yorika et al. [Yorika, 2004] showed by means of actual testing and finite element analyses that the tip form is more important than the probe loading to avoid the mechanical issues in the structure. On the contrary, Martens et al. [Martens, 2004] and Wang et al. [Wang, 2007] demonstrated that the tip shape is not a significant factor for the scrub performance. Both studies agreed on the necessity to have a very high contact precision during the sort.

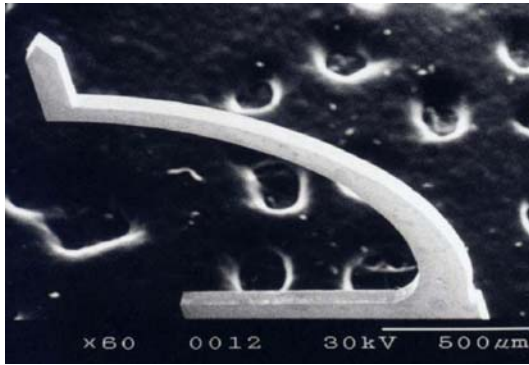


Figure I.20. Microspring probe [Yorika, 2004].

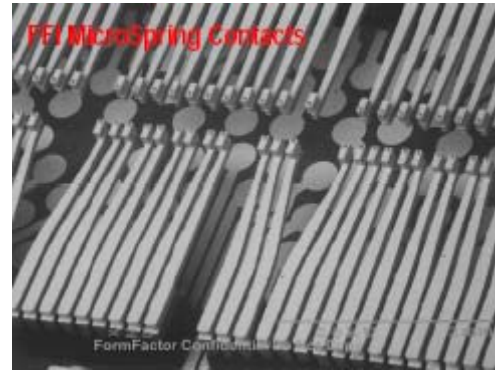


Figure I.21. Microspring probes [Wang, 2007].

II.A.4. Electrical measurements and cleaning

The main purpose of EWS is to check the electrical functionality of the die. To maintain the accuracy of the electrical measurements, the resistance due to the contact has to remain as low and constant as possible. A high contact resistance is likely to induce additional passes to check the functionality of the die and as a consequence additional stress in the pad structure. The contact resistance between the probe needle and the pad is defined according to an equation from Holm [Holm, 1967]:

$$C_{res} = \frac{\rho_{probe} + \rho_{aluminium}}{4} \sqrt{\frac{\pi H}{F}} + \frac{\sigma_{film} H}{F} + R_{trace} \quad (I. 1)$$

where C_{res} is the contact resistance, ρ_{probe} and $\rho_{aluminium}$ are the bulk resistivity of the probe and the aluminium, H is the hardness of the softer material, σ_{film} is the film resistivity, F is the probe force and R_{trace} is the trace resistance contribution.

Several studies have evaluated the contact resistance [Maekawa, 2000; Broz, 1999; Broz, 2006; Liu, 2007], and its variations with parameters such as the over drive (Figure I.22) and the number of touchdowns performed by the probe card (Figure I.23).

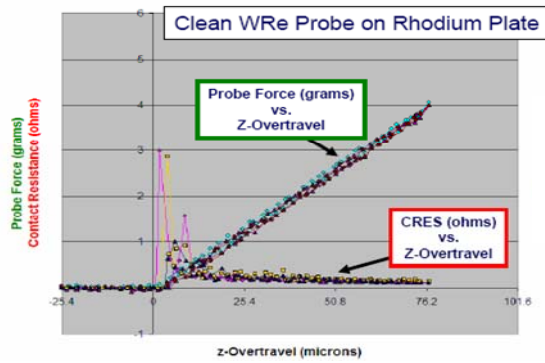


Figure I.22. Probe force and contact resistance versus the applied overdrive [Broz, 2006].

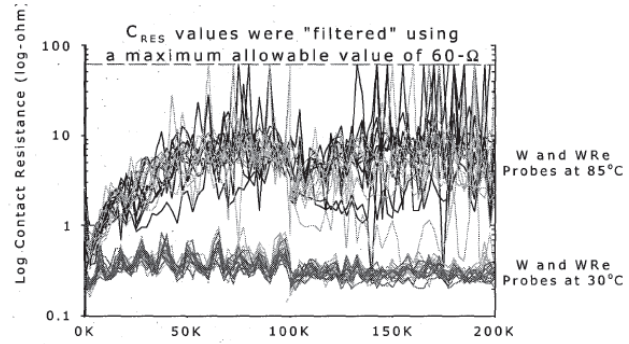


Figure I.23. Cres for W and WRe probes at 30°C and 85°C as a function of the number of touchdowns. A cleaning is performed after 100,000 tests [Broz, 1999].

The studies have underlined the necessity of a regular cleaning of the probe to maintain a low and constant contact resistance (Figure I.23). This cleaning is used to remove particles and other contaminants (e.g. process related organics, bond pad surface oxides...) deposited. This particle contamination causes the contact resistance to increase. The cleaning is made with an abrasive material (a polymer foam containing polishing particles, and coated with a debris collection material to avoid redeposition on pad).

II.B. Failure analysis: pad integrity

Due to applied forces on pads, EWS can induce failures in the oxide layers beneath, requiring failure analysis during the development stages. Nowadays, no non-destructive characterization technique is available to perform large scale study (i.e. on several hundreds of pads) after probing. Thus, the use of destructive method is required to check if mechanical failures had occurred during the test. Among, these techniques, the most common consists in removing the layers on top of the pad and reveal cracks with chemicals. As a result, mechanical failures are observed directly with an optical microscope.

The process to observe the cracks is described by the Figure I.24. Once pads are probed, the alucap is selectively etched by chemicals¹. The etching solution contains acetic acid (2.5% to 10%), nitric acid (<2.5%) and orthophosphoric acid (50% to 97.5%). The latest component reacts with aluminium with an activation temperature of 45°C. Then, the thin TaN/Ta layer is plasma etched. It is a two steps process. The first one is an etching with argon to remove the remaining aluminium and impurities. For the second one, a gas containing CF₄ and O₂ is used to selectively etch the barrier layer since both N and Ta react with fluorine. The plasma parameters have to be fine tuned to etch selectively the layers and obtain a pad free from aluminium and tantalum². Finally, to ease the observation of the cracks under the optical microscope, a chemical attack is performed. The solution used for the aluminium removal is efficient, since the nitric acid infiltrates through the crack to etch the copper and reveal the cracks.

¹ The whole pad integrity test is set up and used by the ST Crolles physical characterization team.

² The 2 steps plasma etching of TaN/Ta has been set up and fine tuned by A. Pofelski and V. Guyader [Pofelski, 2009] from the ST Crolles physical characterization team on the probed dies further presented.

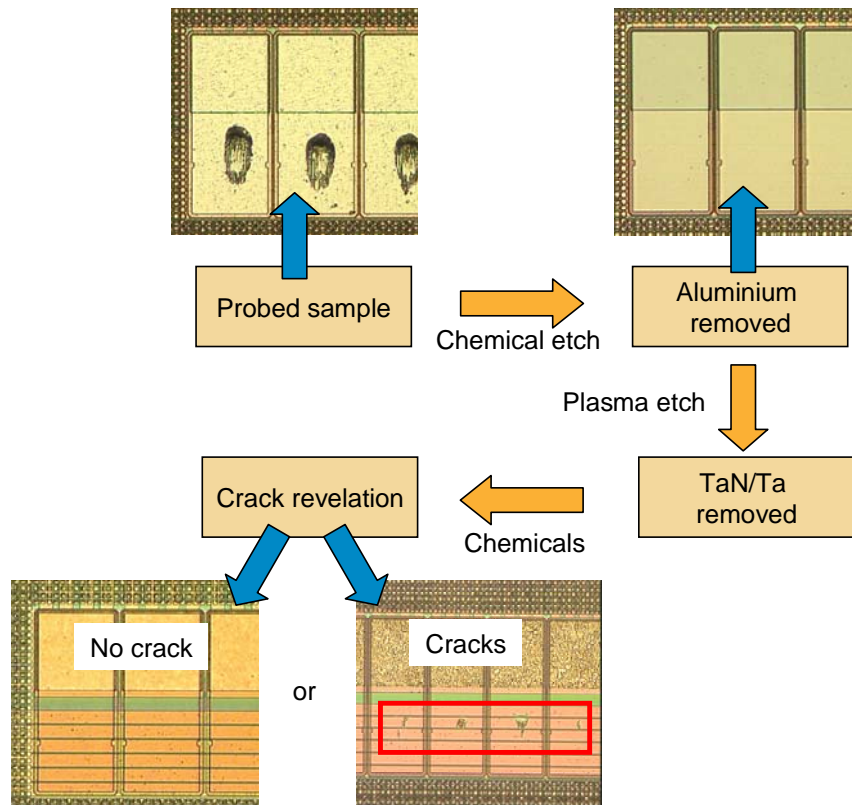


Figure I.24. Successive steps of the pad integrity test for crack revelation.

II.C. Introduction to mechanical issues during probing

In this section, the mechanical damages observed during the EWS are only introduced, since complex mechanical issues are raised by this test and will be further more widely developed.

During the probing test, a scrub mark, whose length, width and depth depends on the test conditions, results from the contact between the hard material of the probe and the ductile aluminium top material of the pad. Thus, each time a pass is made; some aluminium is shifted from its initial location and piled up in front of the needle (see Figure I.25 (b)). While pad integrity tests are performed on harshly probed pads, some cracks appear in the oxide layers underneath the aluminium leading to short circuits and finally to a reject of the die, as shown in Figure I.26. In a more general manner, probing initiate damages and sometimes cracks in the pad structure that reduce the production yield.

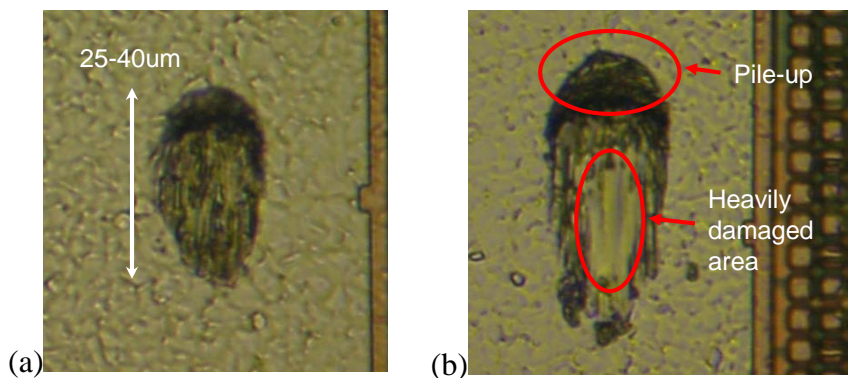


Figure I.25. Probe mark after probing with standard conditions (a) and with harsh conditions (b).

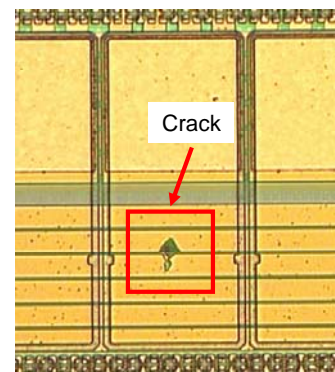


Figure I.26. Crack observed after pad integrity test.

On some structures, and due to the actual trend to reduce the pad pitch, while EWS and wire bonding (or other assemblies) are performed at the same location, probing conditions can induce high damages on the aluminium layer and affect the bondability of wires [Hotchkiss, 2001]. Mercado et al. demonstrated that harsh probing conditions are related to lower shear stresses after ball bonding [Mercado, 2000]. In that frame, Yao et al. [Yao, 2004], found a threshold of 20% for the area of the pad damaged after probing. Similarly, Sauter et al. [Sauter, 2003] has shown that the percentage of probe mark area deeper than $0.6\mu\text{m}$ must be limited to 25-30% of the ball bond area to achieve reliable wire bonding. Finally, EWS is responsible for both failures and damages during testing and issues like non sticking bonding during the assembly steps.

III. Wire bonding and related testing

Wire bonding is the most common way to connect electrically the chip to the package. The other existing ways are controlled collapse chip connection (C4) also called Flip Chip (FC) and the use of through silicon vias, which will be more widely developed section IV. This assembly is part of the back-end (see Figure I.3). Sawed dies are linked to the die attach with the help of wires.

III.A. Test description

Three major techniques were successively used to achieve bonding in microelectronics [Harman, 1997]: thermocompression, ultrasonic and thermosonic bonding. For all these techniques, the wire material is welded. The difference is based on the mean to weld the wire: using heat, ultrasonic energy or both. The bonding on the pad side can be completed either with wedge bonds or more often with ball bonds (see Figure I.27).

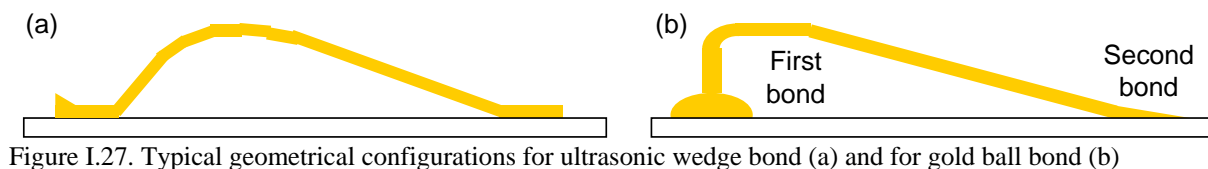


Figure I.27. Typical geometrical configurations for ultrasonic wedge bond (a) and for gold ball bond (b)

III.A.1. First techniques to bond wires

III.A.1.a) Thermocompression bonding

The thermocompression, first bonding technique, was developed by the Bell Laboratories in 1957 [Anderson, 1957; Christensen, 1958] for microelectronics. This method was used only for a few years until ultrasonic wedge bonding replaced it. Thermocompression bonding is a type of solid-phase welding that combines heat and compressive force to plastically deform the welded elements. Surface contaminants (oxides, halogens, sulphur, carbonaceous...) are swept aside and a contact between cleaned surfaces is established. The metallic welded bond is energetically activated, following an Arrhenius relationship. Hence, to achieve such a bonding, the entire device needs to be heated around 300°C . As a consequence, this technique raises issues due to the high bonding temperatures that can damage the entire die and the packaging material.

III.A.1.b) Ultrasonic wedge bonding

Ultrasonic wedge bonding was introduced to the microelectronics industry in the 1960s, and rapidly replaced thermocompression in the device production. Ultrasonic wedge bonding is performed at room temperature, and the welded elements are formed by applying ultrasonic energy through a resonating capillary, while applying a clamping force. The wedge bonding system has to be oriented in a straight line from the first to the second bond before the first one is made. This alignment slows down the entire bonding process by more than 50% compared to the ball bonding.

III.A.2. Thermosonic ball bonding

Ultrasonic power and heat have been first combined by Coucoulas [Coucoulas, 1970]. Today, this is the most common way to perform wire bonding. The combination of ultrasonic and thermocompression welding uses advantages of both techniques. The interface temperature can be lowered to a range between 125°C and 220°C. The ultrasonic energy helps to remove the contaminants and to mature the weld. Moreover, after the ball bond is made, the loop can be formed in any direction that allows fast bonding.

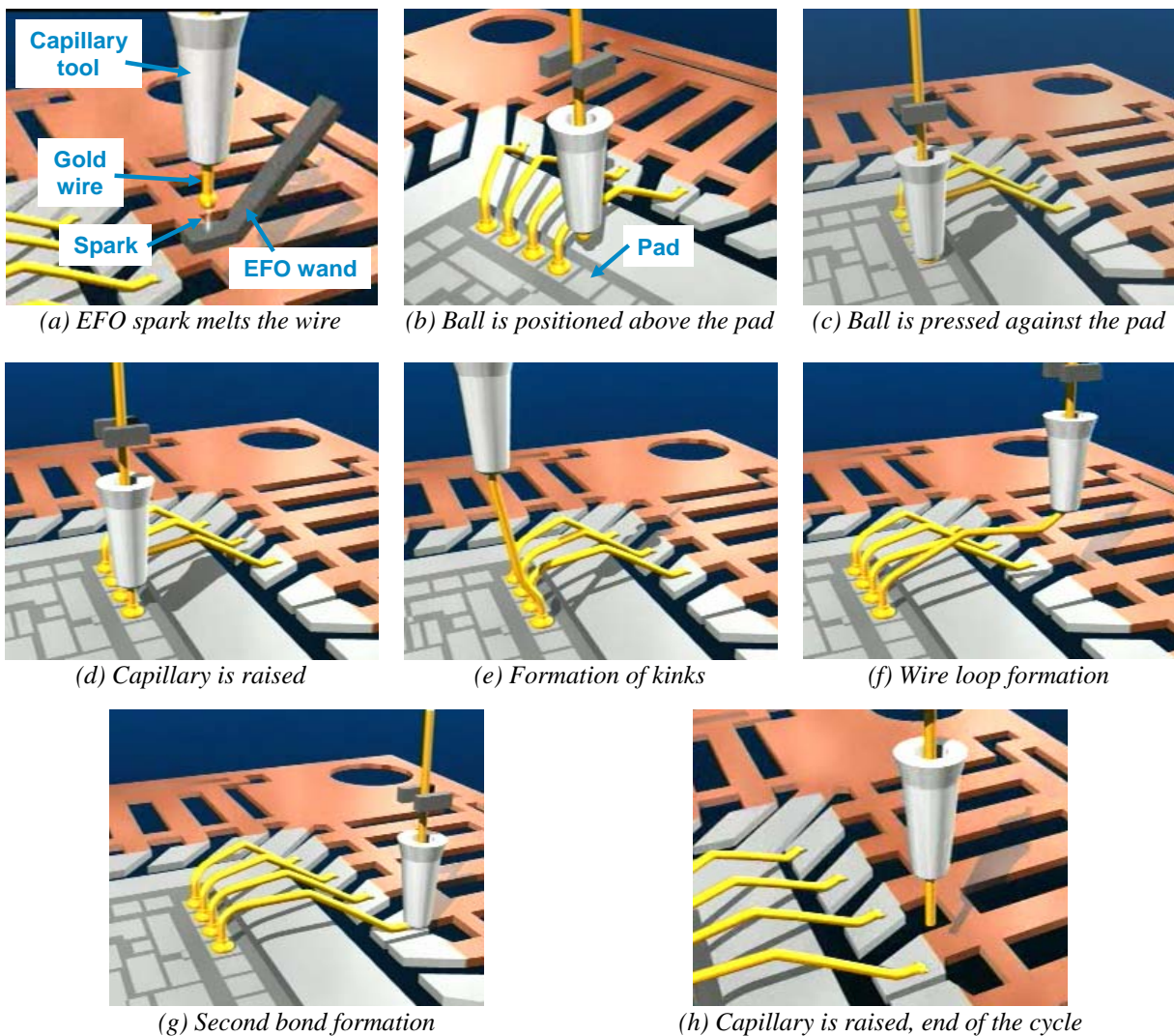


Figure I.28. Simplified procedure for making a ball bonding using the thermosonic method.

The procedure for making a ball-stitch wire connection with a capillary tool is shown in Figure I.28. Only the major stages of the bonding cycle are described. Firstly, the wire is fed

through the capillary, and an Electrical Flame Off (EFO) spark melts the wire. Due to the surface tension of the molten gold, a ball forms at the end of the wire (a). The solidified ball is also known as Free Air Ball (FAB). The wire is pulled into the chamfer so that the wire is positioned against the bottom of the capillary. At the same time, the ball is centred above the pad (b). The tool is lowered to the bond pad and the ball is pressed against it (c). The die is heated all along the wire bonding process. Once the contact between the ball and the pad is established, force and ultrasonic power are applied during a given time and the bond is formed. Then, the capillary rises to the loop height position, leaving the welded ball on the pad (d). Kinks are formed to give to the bond the final desired shape (e) and the bond loop is made as the capillary moves toward the second bond (f). The second bond is also known as stitch or crescent bond. During the formation of the second bond, the wire is deformed against the lead while force, heat and ultrasonic energy are applied (g). After the second bond is made, the capillary tool moves upward, the wire is clamped and the bonding cycle can be repeated (h).

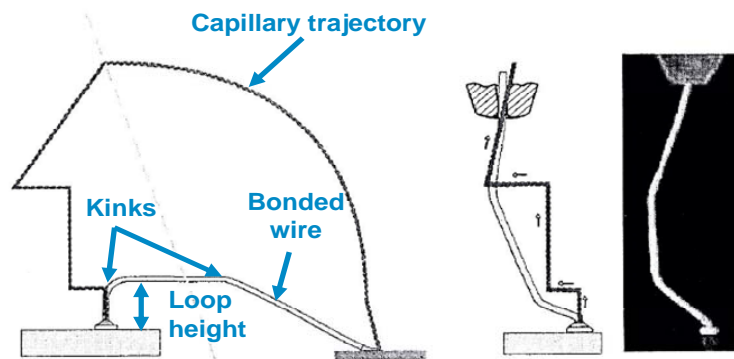


Figure I.29. Capillary trajectory to form the loop shape.

The trajectory of the capillary during the looping stage is one of the key parameter during wire bonding. The trajectory, as depicted by the Figure I.29, has to take into account the distance between the pad and the leadframe, but also the final height of the package; this is achieved by fixing the loop height. Kinks are used to control the looping of the wire, but also to avoid highly elongated areas. Several studies are led to optimize and find out key parameters in the looping behaviour [Saraswati, 2004] or to model and reduce the stress in the wire [Park, 2009].

As described, numerous parameters have to be taken into account for the thermosonic bonding. The heating temperature, the bonding power, the bonding force and the bonding time have an influence on the intermetallic compounds formation (see next section). The bonding time corresponds to the duration of application of the ultrasonic power and the compressive force. The geometrical aspects, such as the loop height, number of kinks and distance between pad and leadframe must also be considered depending on the packaging requirements. Additionally, the impact of the ball on the pad is controlled by the contact force and the capillary velocity. Hence, to achieve a good bonding, all these parameters have to be fine tuned.

III.A.3. Materials in wire bonding

The bonding wire is specified by its electrical, mechanical, thermal and economic properties and considerations. Thus, the mechanical specifications generally provided for the wire are its elongation prior to failure and breaking load. Nowadays, the production units often employ gold wires with smaller and smaller diameters. Yet, a recent trend to move

toward copper emerges. The material considerations, but also advantages and drawbacks of both materials for wire bonding are hereafter described.

III.A.3.a) Gold ball bonding

Since the thermosonic bonding technique is used, gold is selected as bonding material. It has several advantages such as a very good bondability with the aluminium of the pad, good electrical properties and it is inert in air. Due to the use of auto-bonders and to prevent break or stretching when the wire is rapidly pulled through the capillary, strong Au wires with improved breaking loads are required. This is achieved by doping the Au wires with beryllium (5 to 8 ppm) or calcium (5 to 7 ppm). Nevertheless, the dopant concentration is kept small and almost pure gold wires are selected for bonding applications¹. The wire purity is important in the bonding process since it affects the mechanical properties of the wire and as a consequence its reliability.

Gold wires for ball bonding are supplied in the annealed condition. If the wire is hard (as-drawn), the portion immediately above the ball is annealed during its formation to become the weakest part of the wire. This zone is often referred as the Heat Affected Zone (HAZ). The hardness of this zone is lower than the original wire as depicted by Figure I.30 [Hsu, 2007]. The Figure I.31 shows the grains in the wire. The grain structure in the HAZ, after the ball formation, has an intermediate size between the coarse grains of the ball and the fine grains of the thermally stable zone, explaining the reduced mechanical properties of the zone [Wulff, 2003; Dresbach, 2009; Park, 2009].

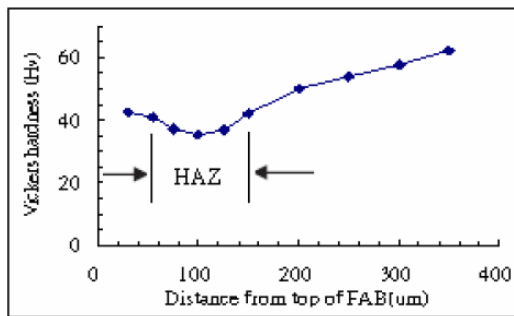


Figure I.30. Micro-Vickers hardness of the FAB, HAZ and as-drawn Au wire [Hsu, 2007].

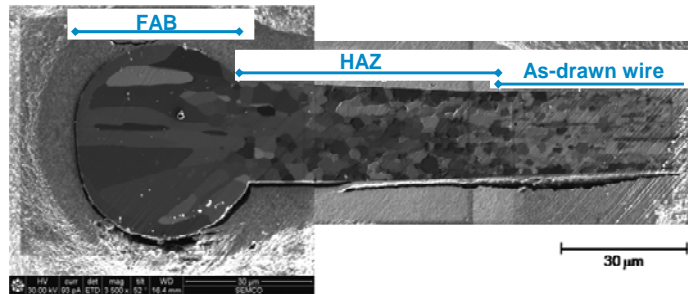


Figure I.31. FIB image of the cross section of a wire with FAB [Park, 2009].

At the interface between the gold ball and the aluminium of the pad, some InterMetallic Compounds (IMC) are formed during the bonding. They are related to the reciprocal diffusion of atoms from one of the material to the other. The IMC coverage during bonding is an important parameter because these compounds ensure a good bond. As shown in the phase diagram, five intermetallic compounds are found in the Au-Al system (Figure I.32). Some properties of the intermetallic compounds are given in Table I.2. After ball bonding, semiconductor chips undergo further processing steps such as encapsulation and curing. The elevated temperatures reached during these steps cause the intermetallic growth (Figure I.33), through atomic diffusion. Many studies evaluated the existing phases, but also the mechanical and electrical properties of these intermetallic compounds [Harman, 1997; Breach, 2004; Breach, 2010; Ji, 2007; Seng, 2007]. They are related to many issues such as reduced electric current conduction and voids (see Figure I.34), which can lead to wire debonding. The reason of these voids is still unclear; in most of the studies, the differences in rate diffusion between gold and aluminium (Al diffuses faster than Au) induces these Kirkendall voids [Harmann, 1997]. However a recent study suggests that these voids are due to the stresses created during

¹ 99% (2N), 99.9% (3N) and 99.99% (4N) pure Au wires are widely used.

the interdiffusion. Since the compounds grow and have various specific volumes, stress gradients can modify the diffusion to create voids [Breach, 2010].

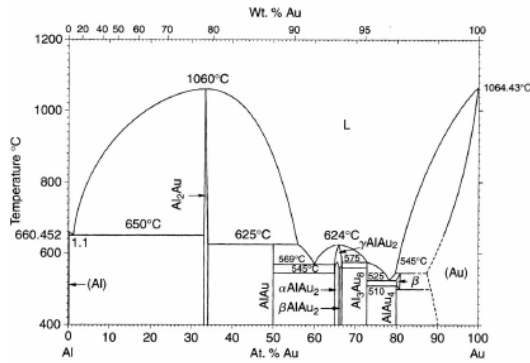


Figure I.32. Aluminium-gold phase diagram [Breach, 2010].

Phase	Composition (at. % Au)	colour	Hardness (GPa)
AuAl ₂	33%	purple	2.58
AuAl	50%	white	2.44
Au ₂ Al	66%	tan	1.28
Au ₈ Al ₃	72%	tan	2.66
Au ₄ Al	80%	tan	3.28

Table I.2. Some properties of the various Au-Al intermetallic compounds [Massalski, 1986].

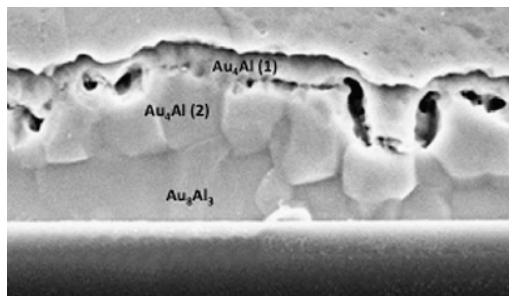


Figure I.33. SEM image of the Au-Al interface after 200h aging [Breach, 2004].

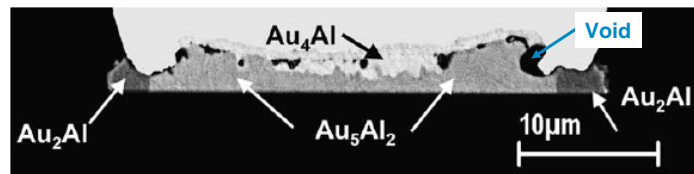


Figure I.34. SEM image of an Au ball bond after 200h aging [Breach, 2004].

III.A.3.b) Copper ball bonding

For cost reduction purposes, the trend is moving toward copper bonding. A comparison of the two wire materials is drawn in the next section. However, among the drawbacks of the copper, it easily oxidizes in air. Oxidation is particularly critical during the formation of the FAB, since this leads to non bonded ball and interconnection damages. Thus, many studies such as Uno et al. [Uno, 2009] investigated various inert gases compositions to achieve a good bonding. Using a forming gas, made from a mixture of 5% H₂ and 95% N₂ is found to ensure a good bondability of the copper wire. Novel options such as coating a thin layer of palladium on the copper wires has also been developed providing promising reliability results [Vath, 2009]. As for gold, the weakest area of the wire is the HAZ. Additionally, the ball is strain hardened during the bonding [Srikanth, 2004].

Similarly to gold bonding, IMC are formed through the same nucleation mechanisms. The phase diagram (Figure I.35) shows that five intermetallic compounds can be found in a Cu-Al system. Some of the properties of the intermetallic compounds are given in Table I.3. Many studies [Seng, 2007; Vath, 2009; Breach, 2010] have been carried out to evaluate the growth of IMC in the Cu-Al system.

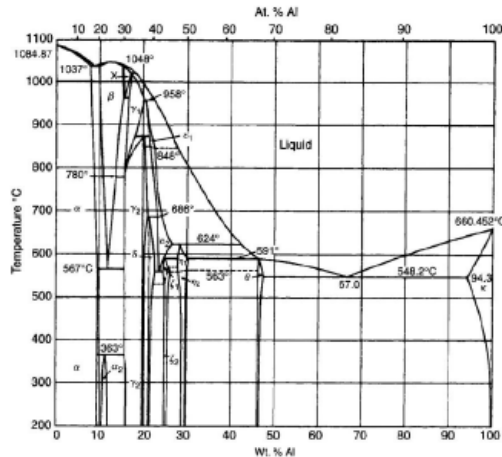


Figure I.35. Aluminium-copper phase diagram [Breach, 2010].

Phase	Composition (at. % Cu)	colour	Hardness (GPa)
CuAl ₂	33%	white-yellow	3.18
CuAl	50%	gray	6.16
Cu ₄ Al ₃	56%	white-yellow	6.04
Cu ₃ Al ₂	60%	white-gray	5.47
Cu ₉ Al ₄	65%	white-light yellow	5.38

Table I.3. Some properties of the various Cu-Al intermetallic compounds [Wulff, 2004].

III.A.3.c) Mechanical and electrical considerations between gold and copper wires

As aforementioned, copper is cheaper than gold. Copper also has other advantages such as a higher electrical conductivity and a higher fusing current. As shown in the Figure I.36, the rate of IMC formation is slower for copper bonding. The slowness of volume diffusion for the Al-Cu IMC is attributed to the reduced connectivity of the lattices of these IMC and the availability of vacancies to transport the atoms. Thus, lower IMC are encountered during copper bonding and no voids are found. Moreover, Cu-Al intermetallics have lower resistivity so the electrical conduction is less altered [Breach, 2010].

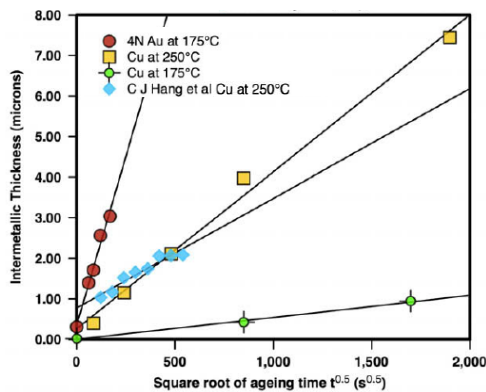


Figure I.36. Total intermetallic thickness of Au-Al and Cu-Al plotted against $t^{1/2}$ [Breach, 2010].

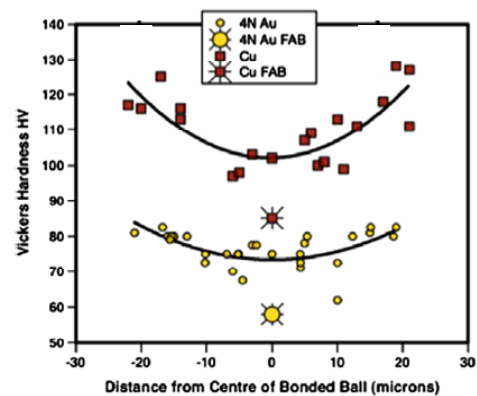


Figure I.37. Variation of the Vickers hardness of bonded Au and Cu balls across the ball diameter [Breach, 2010].

But copper also has several drawbacks that delay its industrial application to wire bonding. The hardness of copper wires is higher than the one of the gold wires as depicted in the Figure I.37. Due to the higher mechanical properties of the copper wire, additional force and power are required to bond the ball, which lead to higher pad deformation and in some cases to damages in the interconnections (cracks in oxide layers). Moreover, the ease of copper to be oxidized in air implies a necessity to use coated wires and/or a forming gas to ensure a reliable bonding.

III.B. Bond characterization tests

In order to check the reliability of wire bonds, several destructive and non destructive tests can be done. The most common methods for evaluating the bonds are the shear test, the wire pull test, and the thermal aging test, which are further described. Others such as chemical opening of the package, etching to obtain the IMC coverage are less easily set up at a large scale. The hereafter described tests are further used during this work to evaluate the bond and the robustness of the structures underneath the pad.

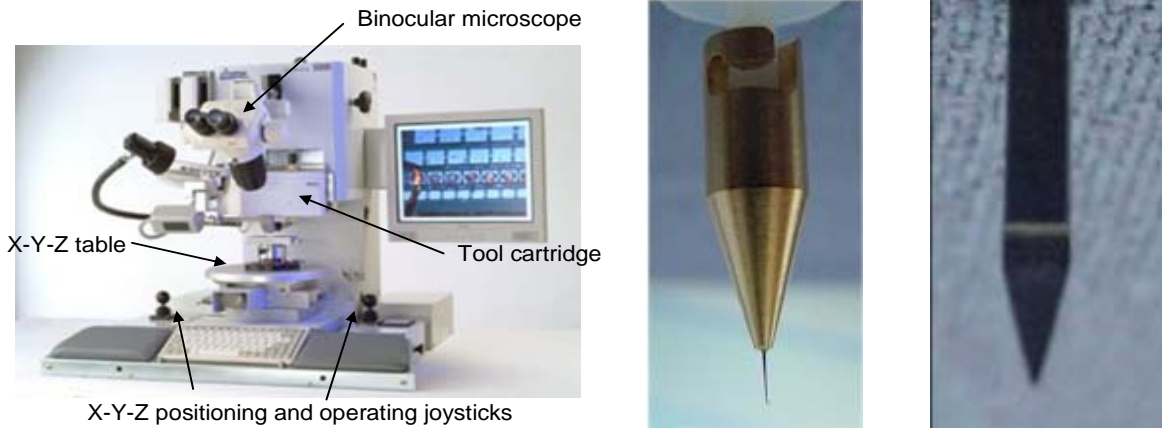


Figure I.38. Pull and shear test apparatus Dage® 5000 (left), hook for wire pull test (centre) and shear tool (right) [Internet 1].

The apparatus depicted in Figure I.38 is currently one of the most common for wire pull and shear tests¹. The sample is placed in horizontal position and viewed from above with an optical microscope. This apparatus allows the control of most of the parameters that can affect measurements during these tests.

III.B.1. Shear test

III.B.1.a) Introduction and apparatus

The shear test is the best way to evaluate the quality of the ball-bond interface, and, thus, allows setting up a ball bonding machine. This test has been introduced in 1966 [Gill, 1966; Arleth, 1967] but it has only been set up for bond strength evaluation by Jellison [Jellison, 1975; Jellison, 1977] after the design of precision shear testers. The principle of the test is to bring a shear tool to the side of a bonded ball, to apply a sufficient force to push it off, and to record the force during the measurement. The test is illustrated in Figure I.39.

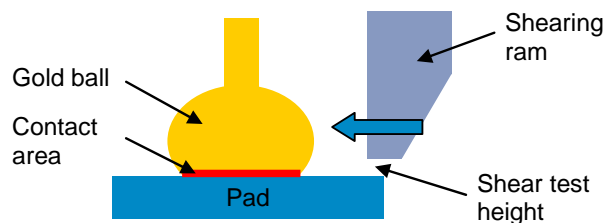


Figure I.39. Schematic drawing of the ball shear test.

The input parameters for the test are the velocity of the shear ram, and the shear tests height. The former is fixed to 250μm/s, and seems to have little effect on the measurement

¹ This apparatus, available at ST Grenoble, is used during further work.

and the latter depends on the sample, and is fixed to $3\mu\text{m}$. The effects of these two parameters are investigated during further experiments.

III.B.1.b) Typical failure modes during Ball shear tests

During the shear test of the ball, the failure modes that can be observed after testing are shown in Figure I.40, after Charles et al. [Charles, 1984]. While performing ball shear test, the failure mode 2 is expected (c). For this mode, some material from the ball remains on the pad, showing a strong bond and thus a good IMC formation during the wire bonding. Weak bond leads to failure mode 1, which corresponds to the absence of wire material remaining on the pad (b). The failure mode number 3 is the most critical since it can be related to a weak structure and can induce reliability issues during the BE operations (d). Failure modes 4 and 5 (e and f respectively) are more related to wrong experimental conditions, and could be easily avoided by using adequate ones. For the former, the shear height is fixed too high, and the shear test is performed close to the top of the ball. For the latter the shear height is too low, and the shear test cannot be successfully achieved.

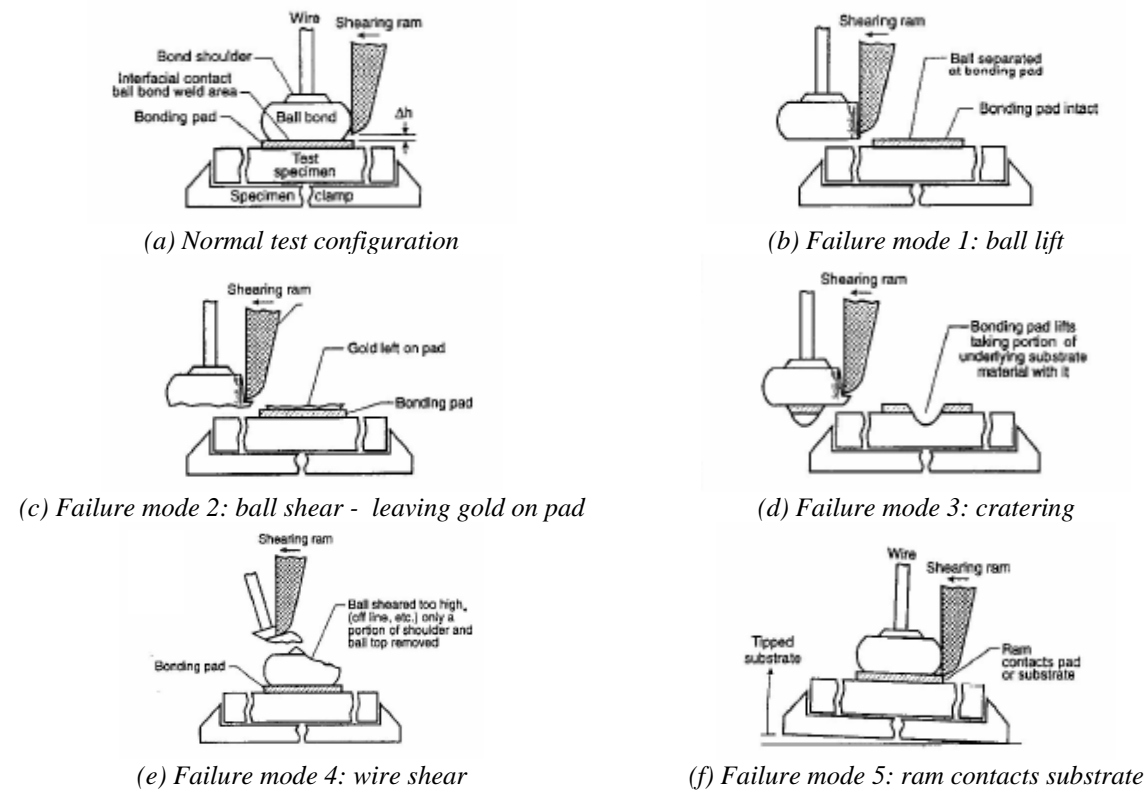


Figure I.40. Normal test configuration and typical failure modes of the ball shear test (after [Charles, 1984]).

The accuracy of the shear test measurement can be reduced by some interference [Harman, 1997]. Firstly, the shear tool must not drag on the substrate. This is related to improper vertical positioning of the tool. Secondly, the shear tool has to be clean in order to keep the shear height constant for successive measurements. Thirdly, the shear tools wear during long use and/or mishandling, leading to deformed surfaces that can also further reduce the test reproducibility.

The output of the test is the measured shear force (SF) prior to failure. Nevertheless, the recorded force depends on the contact diameter between the ball and the pad. Thus, the shear strength (SS) is often introduced to compare the bonding resistance for various wire diameters, or at least for various ball sizes. Two measurements of ball bonded diameter made at right angle are sufficient to obtain the shear strength:

$$SS = \frac{SF}{\frac{\pi}{4} * \left(\frac{d_x + d_y}{2} \right)^2} \quad (\text{I. 2})$$

where d_x and d_y are the ball diameters respectively in the x and y directions.

When the ball is almost perfectly round, a single measurement of the diameter is necessary:

$$SS = \frac{4 * SF}{\pi * D^2} \quad (\text{I. 3})$$

In order to be even more accurate, the measurement of the diameter can be done on the welded area (IMC coverage), that can be achieved by etching the bond pad under the ball with some chemicals that does not attack the ball such as KOH.

III.B.2. Wire bond pull test

III.B.2.a) Introduction and apparatus

The wire bond pull test is used to control the quality of the wire bonding operation. It has been introduced to evaluate the strength of the wire bonds in the 1960s. During a pull test, a hook is positioned below the wire and a vertical tensile load is applied to the wire. Hence, the pull test can be either non-destructive (NDPT) or destructive. For the former, a given low force is applied to the wire, assuming no or little plastic deformation of the wire occurs during the test. The NDPT is used to check the bond strength, only for particular application requiring very high reliability and should be applied to all the wires of the die. On the contrary, the destructive pull test is used while setting up the bonding parameters, during advanced development stages, or when reliability issues are faced.

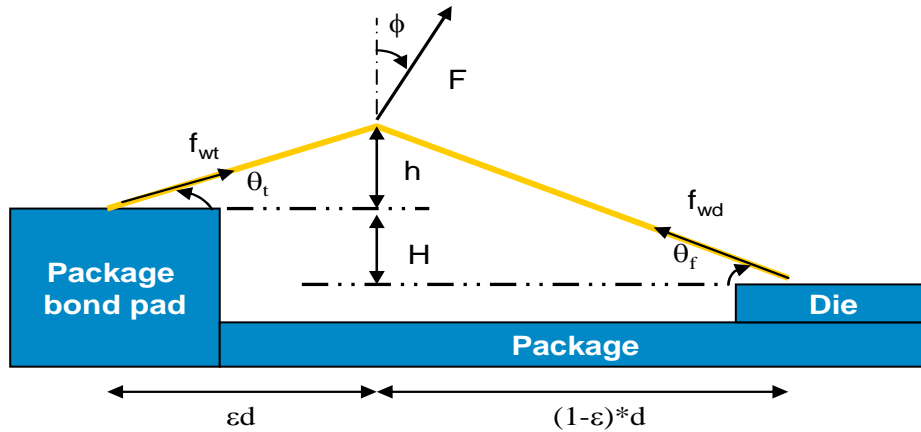


Figure I.41. Geometric variables for wire pull test [Harman, 1978].

The geometrical configuration and the equations defining the resolution of forces were described by Harman [Harman, 1978]. The force in each wire, at break, with a specified pull force, F , at the hook is:

$$f_{wt} = F \left[\frac{\sqrt{h^2 + \epsilon^2 d^2} \left((1 - \epsilon) \cos \phi + \frac{h + H}{d} \sin \phi \right)}{h + \epsilon H} \right] \quad (\text{I. 4})$$

$$f_{wd} = F \left[\frac{\sqrt{1 + \frac{(1-\varepsilon)^2 d^2}{(h+H)^2}} (h+H) \left(\varepsilon \cos \phi - \frac{h}{d} \sin \phi \right)}{h + \varepsilon H} \right] \quad (\text{I. 5})$$

In our case, the wires are pulled straight up ($\phi=0$) and the purpose is to evaluate the ball side of the bond (ε tends toward 1), so finally $F \sim f_{wd}$.

III.B.2.b) Typical failure modes during wire pull test

Several failure modes are observed during the wire pull test, as depicted by Figure I.42. A successfully achieved pull test results in a failure at the wire neck. As mentioned in section III.A.3.a), due to annealing, the HAZ is the weakest part of the wire and as a consequence, failure in the wire is expected to occur in that part. For such failures, the measured force is expected to be a little lower than the breaking load of the wire.

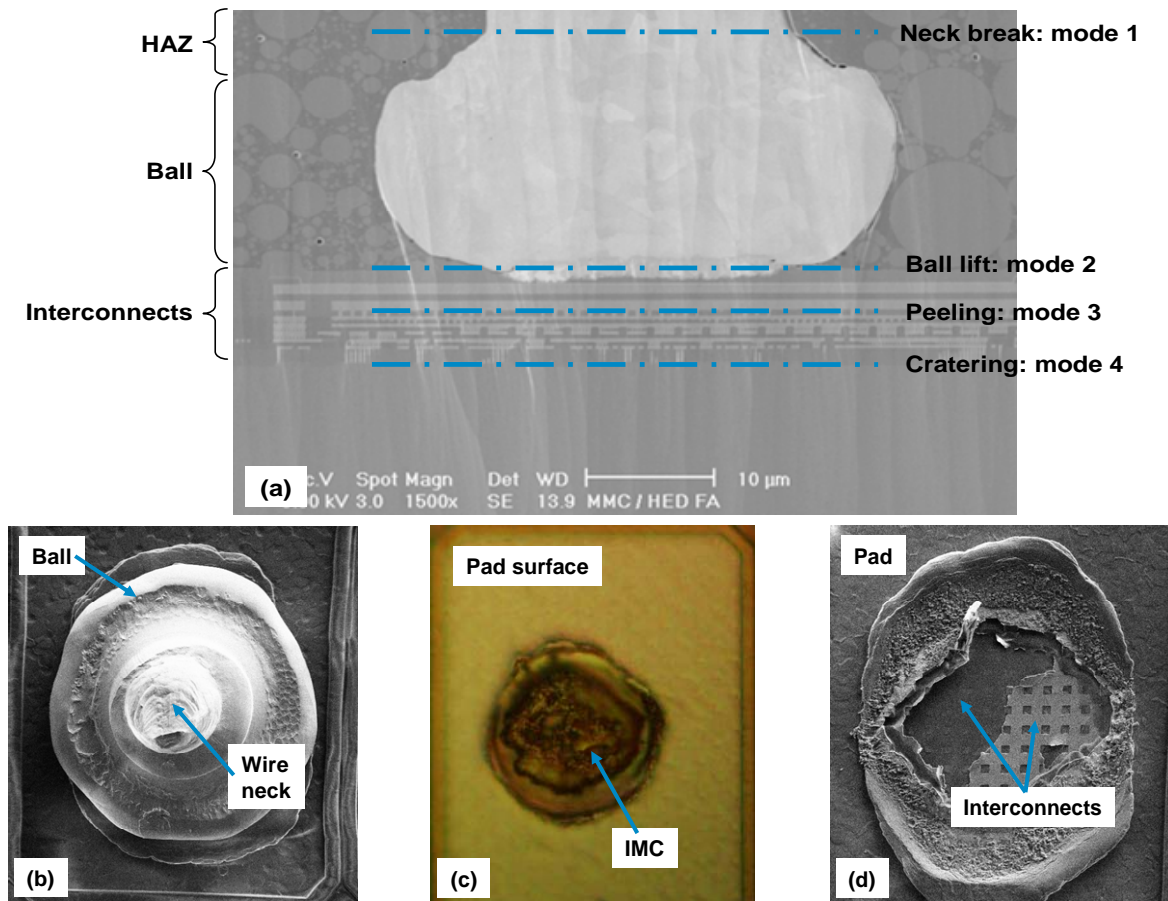


Figure I.42. Typical failure modes during wire pull tests performed close to the ball (a) SEM cross section depicting the 4 failure modes¹, (b) SEM top view of a failure at the wire neck, (c) optical microscopy top view of a mode 2 failure and (d) SEM top view of peeling in the interconnections (mode 3).

Other failure modes are encountered at lower pull forces and can be related either to mechanical weaknesses of the structure or to inadequate bonding conditions. The failure mode

¹ Courtesy of S. Gallois-Garreignot from ST Crolles Mechanical and Thermal Modeling team for the SEM image of the moulded wire.

2 is the ball lift. The ball is detached from the bond pad due to weak welding and thus reduced adhesion of the ball to the pad (b). This is also referred as the Non-Sticking On Pad bond (NSOP). For such failures, some fine tuning of the bonding parameters have to be performed. The peeling or Metal Peel Off (MPO) is a failure within the BEoL, highlighting its weakness for harsh bonding conditions. For such failure, both the HAZ and the bond remain intact (c). The failure mode 4 is the cratering (d). All the FE structure and sometimes additionally some silicon are removed from the die. This is not a very common failure mode, but some paper point out such failures after pull tests [Spaan, 2010].

During wire pull tests, additional failure modes are observed when the force at the second bond cannot be neglected while compared to the pull force on the ball side or when the bond strength of the stitch bond is evaluated. These failures are:

- Midspan failure: the wire fails under the hook location.
- Heel break: the stitch bond fails at the heel.
- Second bond lifting: the second bond is detached from the leadframe, similarly to the NSOP.

III.B.3. Thermal aging

The thermal aging is also known as High Temperature Storage (HTS). During this test, the devices undergo a high temperature (usually 175°C) during a given time (usually 192 hours or 264 hours). Thus, various conditions, such as humidity, thermal heating and aging duration can be investigated. This procedure is dedicated to investigate the effects of the IMC growth. Thermal aging is an accelerated fatigue test, particularly for devices that are submitted to thermal cycling during their lifetimes. The drawbacks of the growth of IMC are increased electrical resistance of the contact and possible failures, which have been developed in III.A.3.

III.C. Mechanical issues during wire bonding

Generally, the mechanical issues induced by wire bonding are found during the development stages when shear and wire pull tests are performed. Though, some issues can be encountered latter during the production. Hence, most of the mechanical issues related to wire bonding are observed after the achievement of all the BE steps. These are only efficiently observed by removing the moulding resin using chemicals to uncover the die. Two main failures are obtained after such opening: broken or lifted wires and bonding pad cratering. The former is related to either heterogeneous thermal mechanical stresses during the assembly steps, non optimized loop profiles or bonding parameters (time, force, heat, ultrasonic power, etc) or contamination of the bonding pad. The driving forces of the latter are the wire bonding parameters, the coefficient of thermal expansion mismatch between aluminium and gold and die stacking issues. During thermosonic bonding, the working temperature can reach 180-200°C. For gold bonding, some bonding degradation can then be related, such as a high increase of the electrical resistance due to the intermetallic growth and possible lift off.

IV. Other processes inducing mechanical failures in pads

The following work mainly focuses on EWS and wire bonding pad related issues. However, many other assembly processes, existing or emerging in the semiconductor industry, lead to similar mechanical failures in the pad structure. In this section, the most common assemblies and their related known failures in interconnections are described.

IV.A. Flip Chip assembly

The flip chip process consists in turning the chip upside-down to directly connect the pad to the substrate. The advantages of this technique are the enhanced thermal package performance and the reduced parasitic impedance provided compared to the wire bonding assembly. Furthermore, the connections over active devices can also be increased. Since this process is much more expensive than wire bonding, its application is still limited to particular products. The two major ways of connecting flipped chips to the pad are described hereafter.

IV.A.1. Solder bump

To achieve the electrical connection between the die and the substrate, solder bumps are used as depicted in Figure I.43. These solder bumps, generally lead free, are commonly made from SnCu, SnAgCu or SnAg alloys. The overall reliability of the process is enhanced by adding a polymer resin, also known as underfill, between the chip and the die.

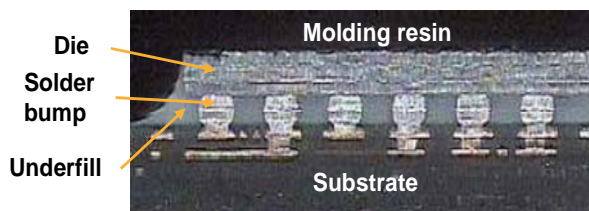


Figure I.43. Schematic side view of a flipped chip assembly [STMicroelectronics].

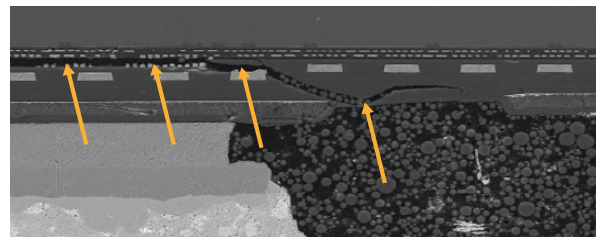


Figure I.44. SEM cross section of sample with failure in the low-k [Chen, 2006]. Delamination is underlined by the arrows.

During this assembly, fractures in interconnections can be observed as shown in Figure I.44 and described by Chen [Chen, 2006]. These failures are often referred as ghost bump, since during X-Ray reliability checking, white dots can be observed under such bumps.

IV.A.2. Copper pillar

This assembly is based on a similar principle to solder bump, by being a recent evolution of the flip chip assembly. A copper pillar is used for the connection instead of the solder bump. On the top of the pillar, a solder joint is added for welding purposes (see Figure I.45). Compared to solder bumps, this assembly enhances thermal and electrical conductions. The density of connections is increased and pillars as small as 60µm can be obtained.

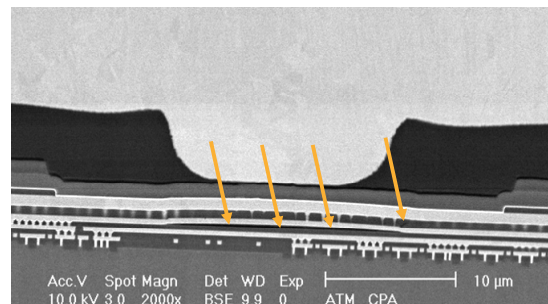
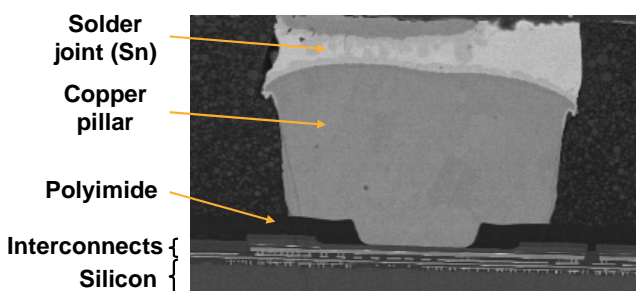


Figure I.45. SEM cross-section views of a flip chip assembly with a copper pillar (left) and zoom at delaminations and failures that can be observed after the assembly, highlighted by the arrows (right) [STMicroelectronics].

The issues met after assembly with copper pillar are similar to those encountered with solder bump: fractures in the interconnections observed as ghost bumps. The Figure I.45 depicts the delamination and failures that can be observed between the layers of the BEOl.

Hence, several numerical studies [Mercado, 2003; Ong, 2009; Zhang, 2010] have been performed to investigate the various design options to reduce the mechanical issues. Among other parameters, the height of the copper pillar, the polyimide (PI) thickness and opening were found to reduce the stress in such structures [Zhang, 2010].

IV.B. Toward 3D integration: TSV

To overcome the Moore's law, the more than Moore law has emerged. Its purpose is to integrate more functionality on devices. It can be achieved by stacking the I.C. in 3 dimensions to reduce the overall size of the dies. The connection between the chip and the substrate is obtained in doing Through Silicon Vias (TSV) that goes through the whole thinned silicon (See Figure I.46). Additionally, micro-bumps, using SnAg alloy, are added for the die interconnection [Gagnard, 2010]. Two main ways of connecting the integrated circuits exist: via-first and via last, depending on the location of the via etching: on the top or at the bottom of the interconnections. This technique is not sufficient enough to connect the dies to the substrate and requires an additional assembly. This is generally achieved by using bumps similarly to flip chip.

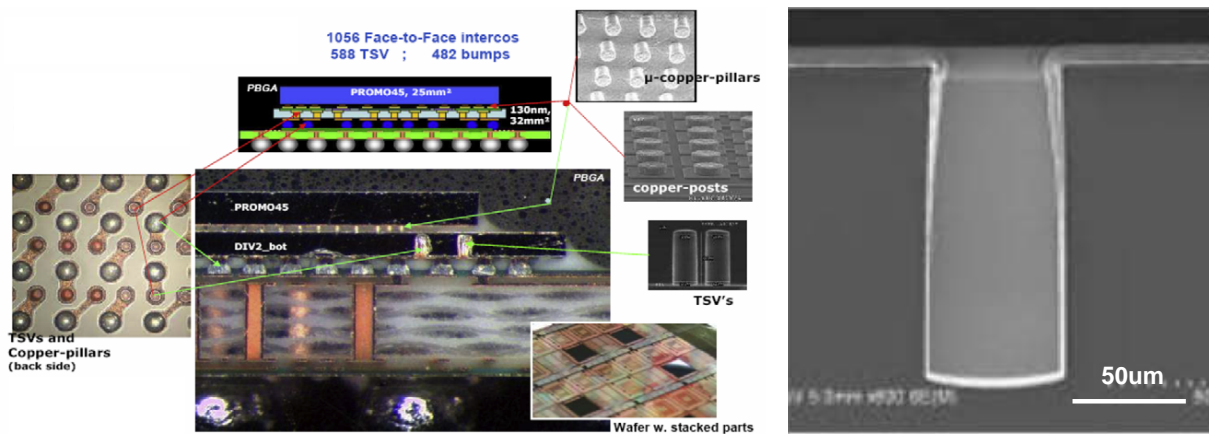


Figure I.46. 3D demonstrator with 45nm and 120nm dies (left) (STMicroelectronics) and cross section of one through silicon via (right) [Gagnard, 2010].

In the literature, most of the mechanical and thermo-mechanical issues are located at geometrical singularities of the TSV, in the lateral barrier of the TSV or at the bottom of the copper via, close to the location of the connection with micro-bumps [Liu, 2009; Khan 2010]. Hence, the mechanical issues in the BEoL for this assembly are up to now limited.

V. Problematic and conclusions

In this first chapter, the main stages of the production from the silicon to the moulded I.C. are swept. These are divided in front end and back end steps, and further described work deals with the interactions related to these parts of the semiconductor industrial process. The pad corresponds to an opening in the integrated circuit to realize the electrical connections of the I.C. with the outer part. With the current trend to reduce the critical dimensions (the pad pitch is particularly a concern in this case) and introduce low-k, brittle dielectric materials, failures are observed in the oxide layers of the interconnections under the pads.

The first source of cracks is the electrical wafer sort, or probing, which is performed to check the electrical functionality of the die at the wafer level. This test consists in establishing a contact between a stiff needle and the top surface of the pad. Thus, during this test, the

aluminium undergoes high damages, and the integrity of the layers underneath is threatened. The probing, its most common parameters and various related test conditions and probe card technologies have been described in this chapter. During the products development stages, the failure analysis, a destructive method, is frequently used as mean to check the pad integrity, by revealing the cracks in the oxide layers. The second source of failures in the BEoL is the various assemblies that enable the electrical connections with the outer part of the integrated circuit. In the semiconductor industry, the most common is the wire bonding, which consists in bonding the pad to the leadframe with gold or copper wires. The bonding conditions applied during the process can induce failures and thus lead to non-functioning dies and yield loss. Some destructive techniques such as the wire pull, the shear and the high thermal storage tests have been developed in the frame of evaluating the reliability of the bonded wires. These tests and the failure modes observed are presented. Hence, in this first chapter, the parametric test and assemblies inducing mechanical issues under the pad during the process flow and the apparatuses used to check the pad integrity are depicted.

In the next chapters, the experimental and numerical tools developed during this work to mimic these operations and reproduce their loading conditions and failures are described. These tools are required to improve the knowledge on the failures locations and mechanisms induced by these processes, since both are widely unknown. On the one hand, the nanoindentation tool, which is generally used to evaluate the mechanical properties of thin films, is applied to the pad and a dedicated methodology for data analysis is developed. On the other hand, a 3D explicit numerical model is set up to reproduce the EWS. Such a complex model is required since the contact between the needle and the probe, the needle dynamics and various pad designs have to be reproduced. Moreover, the issues faced during EWS and assemblies involve many scales from the hundreds of nanometres to few millimetres. Then, these experimental and numerical tools and methodologies are applied to actual test conditions and pad designs, and compared with probing and wire bonding. Some correlations between the analysis techniques, but also some limitations are emphasized.

The pad mechanical robustness is the driving force of the overall study, and the main purpose is thus to provide tools to understand, predict and reduce the fracture hazard in the interconnections. As a consequence, to fit with industrial requirements in fast and reliable tools, some nanoindentation methodologies and rules for pad evaluation and an implicit numerical model are developed to evaluate the robustness of the architectures. The description, uses and limitation of these tools are finally depicted in the frame of providing guidelines for pad design.

References

- [Aigo, 1984] Aigo, S., Unitary probe assembly, US Patent no. 4.480.223, October, 20, 1984.
- [Anderson, 1957] Anderson, O.L., Christensen, H., and Andreatch, P., Technique for connecting electrical leads to semiconductors, Journal of applied physics 28, No. 8, 1957, pp. 923.
- [Arleth, 1967] Arleth, J.A. and Demenus, R.D., A new test for thermocompression microbonds, Electronic products, vol. 9, May 1967, pp. 92-94.
- [Bove, 1974] Bove, R., Probe contactor having buckling beam probes, US patent no. 3.806.801 from April, 23, 1974.
- [Breach, 2004] Breach, C.D. and Wulff, F.W., New observations on intermetallic compound formation in gold ball bond: general growth pattern and identification of two forms of Au₄Al, Microelectronics reliability, vol. 44, 2004, pp. 973-981.
- [Breach, 2010] Breach, C.D. and Wulff, F.W., A brief review of selected aspects of the materials science of ball bonding, Microelectronics reliability, vol. 50, 2010, pp. 1-20.
- [Brillet, 2007] Brillet-Rouxel, H., Etude expérimentale et numérique des phénomènes de fissure dans les interconnexions de la microélectronique, Thesis Science et Génie des matériaux. Grenoble, France, University Joseph Fournier, 2007, 190 p.
- [Broz, 1999] Broz, J.J. and Rincon, R.M., Probe contact resistance variations during elevated temperature wafer test, IEEE International test conference, 1999, pp. 396-405.
- [Broz, 2006] Broz, J.J., Moore, R. and Mui, P., Novel methodologies for assessing on-line probe process parameters, Southwest test workshop, June, 11-14, 2006.
- [Byrnes, 1977] Byrnes, H.P. and Wahl, R., Contact for an electrical contactor assembly, US patent no. 4.027.935 from June, 7, 1977.
- [Charles, 1984] Charles, H.K., Clatterbaugh, G.V. and Weiner, J.A., The ball bond shear test: its methodology and applications, Semiconductor processing, ASTM STP 850, Dinesh C. Gupta, Ed., American society for testing and materials, 1984, pp. 425-457.
- [Chen, 2006] Chen, K.M., Jiang, D.S., Kao, N.H. and Lai, J.S., Effects of underfill materials on the reliability of low-k flip-chip packaging, Microelectronics reliability 46, 2006, pp.155-163.
- [Chin, 2007] Chin, P.-C., Hu, C.-Y., Hsu, H.-C., Fu, S.-L., Yeh, C.-L. and Lai, Y.-S., Characteristic of heat affected zone in thin gold wire and dynamic transient analysis of wire bonding for microstructure of Cu/low-k wafer, International microsystems, packaging, assembly and circuit technology, IMPACT 2007, pp. 297-300.
- [Chiu, 2009] Chiu, J.-T. and Chang, D.-Y., A new probe design combining finite element method and optimization used for vertical probe card in wafer probing, Precision engineering 33, 2009, pp. 395-401.
- [Christensen, 1958] Christensen, H., Electrical contact with thermo-compression bonds, Bell laboratory record, April 1958, pp. 127-130.
- [Coucoulas, 1970] Coucoulas, A., Ultrasonic bonding method, US patent no. 3.507.033 from April, 21, 1970.

[Dresbach, 2009] Dresbach, C., Lorenz, G., Mittag, M., Petzold, M., Milke, E. and Müller, T., Local hardening behaviour of free air balls and heat affected zone of thermosonic wire bond interconnections, European microelectronics and packaging conference, 2009, pp. 1-8.

[Evans, 1988] Evans, A., Multi-level test probe assembly for IC chips, US Patent no. 4.719.417 from January, 12, 1988.

[Gagnard, 2010] Gagnard, X. and Mourier, T., Through silicon via: from the CMOS imager sensor wafer level package to the 3D integration, Microelectronic engineering 87, 2010, pp. 470-476.

[Gill, 1966] Gill, W., and Workman, W., Reliability screening procedures for integrated circuits, fifth annual symposium on the physics of failure in electronics, 1966, pp. 101-142.

[Harman, 1978] Harman, G.G. and Cannon, C.A., The microelectronic wire bond pull test – how to use it, how to abuse it, IEEE transactions on components, hybrids and manufacturing technology, vol. chmt-1, no. 3, September 1978.

[Harman, 1997] Harman, G.G., Wire bonding in microelectronics: materials, processes, reliability and yield. Second edition, Ed: McGraw-Hill, 1997.

[Harmon, 1969] Harmon, R.C., Microelectronic test probe with defect marker access, US Patent no. 3.445.770, May, 20, 1969.

[Hauck, 2009] Hauck, T., Schmadlak, I., Argento, C. and Müller, W.H., Damage risk assessment of under pad structures in vertical wafer probe technology, European Microelectronics and Packaging conference, 2009, pp. 1-5.

[Holm, 1967] Holm, R., Electric contacts, Theory and applications, 4th edition, New York, Springer, 1967.

[Hotchkiss, 2001] Hotchkiss, G., Ryan, G., Subido, W., Broz, J., Mitchell, S., Rincon, R., Rolda, R., and Guimbaolibot, L., Effects of probe damage on wire bond integrity, IEEE Electronics components and technology conference, 2001.

[Hsu, 2007] Hsu, H.C., Chang, W.Y., Fu, S.L., Yeh, C.L. and Lai, Y.S., Dynamic finite element analysis on underlay microstructure of cu/low-k wafer during bonding process, International conference on electronic materials and packaging, 2007, pp. 1-6.

[Hung, 2009] Huang, Y., Hsu, T., Liu, G., Hung, W., Lin, C. and Yang, D., An experimental work and analysis of vertical cobra probing on low-k wafer, IEEE Semiconductor wafer test workshop 2009, June 7-10.

[Internet 1] www.nordson.com, consulted on 09/28/2010.

[Jellison, 1975] Jellison, J.L., Effect of the surface contamination on the thermocompression bondability of gold, IEEE transactions on parts, hybrids and packaging, vol. php-11, no. 3, September 1975.

[Jellison, 1977] Jellison, J.L., Kinetics of thermocompression bonding to organic contaminated gold surfaces, IEEE transactions on parts, hybrids and packaging, vol. php-13, no. 2, June 1977.

[Ji, 2007] Ji, H., Li, M., Wang, C., Bang, H.S. and Bang, H.S., Comparison of interface evolution of ultrasonic aluminium and gold wire wedge bonds during thermal aging, Material science and engineering A, vol. 447, 2007, pp. 111-118.

[Khan, 2010] Khan, N., Rao, V.S., Lim, S., We, H.S., Lee, V., Zhang, X., Liao, E.B., Nagarajan, R., Chai, T.C., Kripesh, V. and Lau, J.H., Development of 3-D silicon module

with TSV for system in packaging, IEEE Transactions on components and packaging technologies, vol. 33, no. 1, March 2010, pp. 3-9.

[Kim, 2003] Kim, H.-J., Lee, J.Y., Paik, K.-W., Koh, K.-W., Won, J., Choe, S., Lee, J., Moon, J.-T. and Park, Y.-J., Effects of Cu/Al intermetallic compound (IMC) on copper wire and aluminium pad bondability, IEEE transactions on components and packaging technologies, vol. 26, no. 2, June 2003, pp. 367-374.

[Leu, 1995] Leu, J., Holland, S., Monnig, K.A. and Ho, P.S., Evaluation of low dielectric constant materials for on-chip interconnections – an industry/university research collaboration, Proceedings of the eleventh biennial university/government/industry microelectronics symposium, pp. 122-125.

[Liu, 2007] Liu, D.S., Shih, M.K. and Huang, W.H., Measurement and analysis of contact resistance in wafer probe testing, Microelectronics reliability 47, pp. 1086-1094.

[Liu, 2009] Liu, X., Chen, O., Dixit, P., Chatterjee, R., Tummala, R.R. and Sitaraman, S.K., Failure mechanisms and optimum design for electroplated copper through-silicon vias (TSV), IEEE Electronic components and technology conference, 2009, pp. 624-629.

[Maekawa, 2000] Maekawa, S., Takemoto, M., Kashiba, Y., Deguchi, Y., Miki, K. and Nagata, T., Highly reliable probe card for wafer testing, IEEE Proceeding from the Electronic components and technology conference, 2000, pp. 1152-1156.

[Maex, 2003] Maex, K., Baklanov, M.R., Shamiryan, D., Lacopi, F., Brongersma, S.H. and Yanovitskaya, Z.S., Low dielectric constant materials for microelectronics, Journal of applied physics, vol. 93, Issue 11, 2003, pp. 8793-8841.

[Mann, 2004] Mann, W.R., Taber, F.L., Seitzer, P.W. and Broz, J.J., The leading edge of production wafer probe test technology, International test conference 2004, p. 1168-1195

[Martens, 2004] Martens, R. and Levy, L., Optimization of Microspring® contact design parameters for low pressure probing, Southwest test workshop, June, 7, 2004.

[Massalski, 1986] Massalski, T.B., Binary phase diagrams handbook, Ed: ASM, Ohio, 1986.

[Mercado, 2000] Mercado, L., Radke, R., Rutson, M., Tran, T.A., Williams, B., Yong, L., Chen, A., and Chen, S., Reliability of multi-layer aluminium capped copper interconnect structures, IEEE International electronics manufacturing symposium, 2000, pp. 84-93.

[Mercado, 2003] Mercado, L.L., Kuo, S.-M., Goldberg, C. and Frear D., Impact of flip chip packaging on copper/low-k structures, IEEE transactions on advanced packaging, vol. 26, no. 4, November 2003, pp. 433-440.

[Moore, 2000] Moore, T.M., Hartfield, C.D., Anthony, J.M., Ahlburn, B.T., Ho, P.S. and Miller, M.R., in: D. G. Seiler, A. C. Diebold, T. J. Shaffner, R. McDonald, W. M. Bullis, P. J. Smith, E. M. Secula (Eds.), Characterization and metrology for ULSI technology 2000, Gaithersburg, USA, June 26-29, 2000, American institute of physics symposium proceedings 550, 2001, pp. 431-439.

[Ohkubo, 1986] Ohkubo, M. and Yoshimitsu, Y., Complex probe card for testing a semiconductor wafer, US Patent no. 4.567.433 from January, 28, 1984.

[Ong, 2009] Ong, J., Tay, A., Zhang, X., Kripesh, V. and Lim, Y.K., Optimization of the thermomechanical reliability of a 65nm Cu/low-k large-die flip chip package, IEEE transactions on components and packaging technology, vol. 32, no. 4, December 2009, pp. 838-848.

[Padovani, 2001] Padovani, A.M., Rhodes, L., Riester, L., Lohman, G., Tsuie, B., Conner, J., Bidstrup Allen, S.A. and Kohl, P.A., Porous Methylsilsesquioxane for low-k dielectric applications, *Electrochemical and solid-state letters*, 4 (11), pp. F25-F28.

[Park, 2009] Park, J.-W., Yoon, Y.-B. and Lee, J.-G., The fracture resistance of bonded Au wires for interconnection, *IEEE transactions on components and packaging technologies*, vol. 32, no. 1, March 2009, pp. 106-109.

[Pofelski, 2009] Pofelski, A., Guyader, V. and Pignard, S., Développement d'analyses de caractérisation physique et de nouvelles techniques de préparation d'échantillons à base de procédés gravure plasma, INP Grenoble, France, Rapport de stage de fin d'étude, 2009.

[Saraswati, 2004] Saraswati, Theint, E.P.P., Stephan, D., Wulff, F.W., Breach, C.D. and Calpito, D.R.M., Looping behaviour of gold ball bonding wire, *Electronic packaging technology conference, EPTC 2004*, pp. 723-728.

[Sauter, 2003] Sauter, W., Aoki, T., Hisada, T., Miyai, H., Petrarca, K., Beaulieu, F., Allard, S., Power, J. and Agbesi, M., Problems with wire bond on probe marks and possible solutions, *IEEE electronic components and technology conference*, 2003, pp. 1350-1358.

[Seng, 2007] Seng, Y.L., Characterization of intermetallic growth for gold bonding and copper bonding on aluminium metallization in power transistors, *IEEE electronics packaging technology conference*, 2007, pp. 731-736.

[Spaan, 2010] Spaan, E., Ooms, E., van Driel, W.D., Yuan, C.A., Yang, D.G. and Zhang, G.Q., Wire bonding the future: a combined experimental and numerical approach to improve the Cu-bonding wire quality, *IEEE 11th international conference on thermal, mechanical and multiphysics simulation and experiments in micro-electronics and micro-systems, Eurosime 2010*.

[Srikanth, 2004] Srikanth, N., Murali, S., Wong, Y.M. and Vath III, C.J., Critical study of thermosonic copper ball bonding, *Thin solid films*, vol. 462-463, 2004, pp. 339-345.

[Uno, 2009] Uno, T., Terashima, S. and Yamada, T., Surface enhanced copper bonding wire for LSI, *IEEE electronic components and technology conference*, 2009, p. 1486-1495.

[Vallauri, 2007] Vallauri, R., Gervasoni, M., Crippa, R. and Lazzari, S., Route60: a new vertical probing technology, *IEEE Semiconductor wafer test workshop*, June 3-6, 2007.

[Vath, 2009] Vath III, C.J., Gunasekaran, M. and Malliah, R., Factors affecting the long term stability of Cu/Al ball bonds subjected to standard and extended HTS, *IEEE electronic packaging technology conference*, 2009, pp. 374-380.

[Vella, 2003] Vella, J.B., Adhihetty, I.S., Junker, K. and Volinsky, A.A., Mechanical properties and fracture toughness of organo-silicate glass (OSG) low-k dielectric thin films for microelectronic applications, *International journal of fracture* 119/120, p. 487-499

[Volinsky, 2003] Volinsky, A.A., Palacio, M-L.B. and Gerberich, W.W., 'Incompressible' pore effect on the mechanical behaviour of low-k dielectric films, *Material research society symposium proceeding*, vol. 750, pp. 567-572.

[Wang, 2007] Wang, Y. and Leong, A., Contact precision optimization to improve scrub performance, *Southwest test workshop*, June, 7, 2004.

[Wulff, 2003] Wulff, F.W., Breach, C.D. and Dittmer, K., Crystallographic texture of drawn gold bonding wires using electron backscattered diffraction (EBSD), *Journal of material science letters*, vol. 22, 2003, pp. 1373-1376.

[Wulff, 2004] Wulff, F.W, Breach, C.D., Stepahn, D., Saraswati and Dittmer, K.J., Characterisation of intermetallic growth in copper and gold ball bonds on aluminium metallisation, IEEE Electronics packaging technology conference, 2004, pp. 348-353.

[Yao, 2004] Yao, Y.F., Xiong, Z.P., Gu, X. and Chua, K.H., Assembly process development of 50um fine pitch wire bonded devices, IEEE electronic components and technology conference, 2004, pp. 365-371.

[Yorika, 2007] Yorika, J., Haga, T., Hirata, Y. and Shimada, S., Non-damage probing and analysis of ILD damage at scrub marks, Southwest test workshop, June, 7, 2004.

[Zhang, 2010] Zhang, X.R., Zhu, W.H., Liew, B.P. and Gaurav, M., Copper pillar bump structure optimization for flip chip packaging with Cu/low-k stack, IEEE International conference on thermal, mechanical and multiphysics simulation and experiments in micro-electronics and micro-systems, Eurosime 2010, pp. 715-721.

Chapter 2:

Comprehension of the mechanical failures: Presentation of the experimental and numerical tools set up

Outline

I.	Experimental techniques for failure analysis	- 46 -
I.A.	3D observation of the pad surface: Optical profilometry	- 46 -
I.A.1.	Principle	- 46 -
I.A.2.	Expectations and limitations of the technique.....	- 47 -
I.B.	3D reconstruction of the pad volume: FIB/SEM reconstruction	- 47 -
I.B.1.	FIB and SEM principles.....	- 48 -
I.B.2.	Advantages and drawbacks of the technique	- 48 -
I.C.	Additional tools used to evaluate the pad damages	- 49 -
I.C.1.	Electron backscatter diffraction	- 49 -
I.C.2.	Energy dispersive X-Ray	- 49 -
II.	Experimental evaluation of the mechanical robustness of pads: nanoindentation.....	- 50 -
II.A.	Principle and apparatus	- 50 -
II.A.1.	Presentation of the technique.....	- 50 -
II.A.2.	Berkovich and cube corner indenter tips	- 51 -
II.A.3.	General description of the indentation process.....	- 52 -
II.B.	Determination of material properties	- 53 -
II.B.1.	Measurement of hardness and elastic modulus	- 53 -
II.B.2.	Oliver and Pharr model and limitations.....	- 53 -
II.B.3.	Parameters affecting the indentation curves	- 55 -
II.B.4.	Continuous stiffness measurement	- 56 -
II.C.	Considerations for the indentation of multilayer structures.....	- 56 -
II.D.	Nanoindentation experiments for pad robustness evaluation	- 58 -
III.	EWS numerical model	- 58 -
III.A.	Model description	- 59 -
III.A.1.	Geometrical model	- 59 -
III.A.2.	Material Properties	- 60 -
III.A.3.	Loading conditions and contact checking	- 61 -
III.A.4.	Physical and numerical parameters	- 62 -
III.B.	EWS numerical model set up: calibration	- 66 -
III.B.1.	Experiment and calibration for CMOS 120nm pad model.....	- 67 -
III.B.2.	Experiment and calibration for CMOS 45nm pad model.....	- 71 -
III.B.3.	Modal analysis of the probe behaviour.....	- 74 -
IV.	Conclusions	- 77 -
References	- 79 -

In the first chapter, the pad elaboration has been described. The industrial tests and processes to check the electrical functions and connect the die electrically to the outer part of the integrated circuit are known to induce mechanical failures in the interconnections, and more specifically in the oxide layers. These failures are related to leakage and short circuits and must be avoided. Thus, some dedicated experimental and numerical tools and analyses are firstly required to improve the knowledge of the failure mechanisms.

In this second chapter, the experimental techniques (FIB/SEM, optical profilometry...) are firstly presented to observe the crack locations and understand the root causes leading to these fractures. Then, some methodologies are described in order to reproduce the loading conditions and failures observed during the actual tests and assemblies. Hence, the nanoindentation technique and the Oliver and Pharr model used to extract the results during the experiments are presented. The compressive loading applied during the indentation test corresponds to a simplified loading condition compared to EWS or wire bonding. Additionally, an explicit finite element model is set up to mimic the probing. The model with related assumptions and also physical and numerical parameters is described. To obtain a faithful model, it is calibrated with some dedicated experiments. Finally, a modal analysis of the finite element model is performed to check some of dynamical effects related to probing.

I. Experimental techniques for failure analysis

The characterization of surfaces and volumes at the micrometre scale requires adequate tools. In some cases such as scrub mark control after probing, optical microscope is sufficient, but generally, such viewing provides limited information. Hence, the imaging tools used during the following work and their principle and application domains are described. Then, some other techniques coupled to these imaging apparatuses such as electron backscatter diffraction and energy dispersive X-Ray are briefly mentioned.

I.A. 3D observation of the pad surface: Optical profilometry

Optical microscopy profilometry is used to obtain a 3D topography of a surface. The principles, application but also limitations of this technique are described.

I.A.1. Principle

The optical profilometry principle is based on light interferences. There are 2 distinct techniques: Phase Shift Interference (PSI) and Vertical Shift Interference (VSI). The former is dedicated to roughness measurements and the latter, using white light, is based on the principle of vertical sweeping to evaluate the topography of the samples. Hence the VSI mode is suitable to evaluate the pad wear due to testing and assemblies.

A schematic view of the instrument is depicted in Figure II.1. To perform the mapping of a surface, the light from the source goes on a semi transparent mirror and is focused by the lens on the microscope's objective. Then, the light is separated in two beams by a splitter:

- The first one reflects back on a mirror located in the microscope objective. Then, the beam is reflected by the semi reflecting mirror of the beam splitter on the sensor of a camera.
- The second beam goes a first time through the semi reflecting mirror of the splitter to reflect back on the surface of the sample. Then, it goes a second time through the mirror to interfere with the first beam. The sensor of the camera saves the intensity and the position of the interference fringes.

The principle of the measurement with a white light consists in reducing continuously the size between the objective of the microscope and the sample. Therefore, the fringes of

interference go on the sample. Finally, once the measurement is completed, algorithms calculate the heights of the areas of the sample to give a 3 dimensional image.

The interference is constructive while the optical paths of the 2 beams are equals. The light intensity reaches its maximum. While the difference in the optical path is k time $1/2$ (k is an integer), then the interference is destructive and the light intensity observed by the camera is null. Thus, the measurement provides a serial of interferences fringes on the sample depending on the distance between the surface of the sample and the objective.

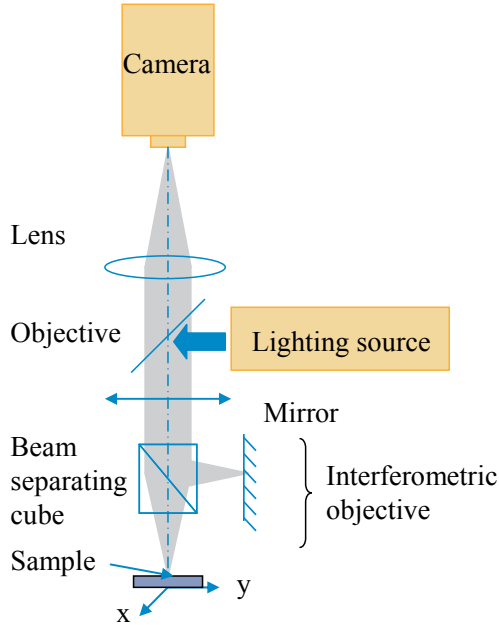


Figure II.1 Schematic principle of the optical microscope profilometry.

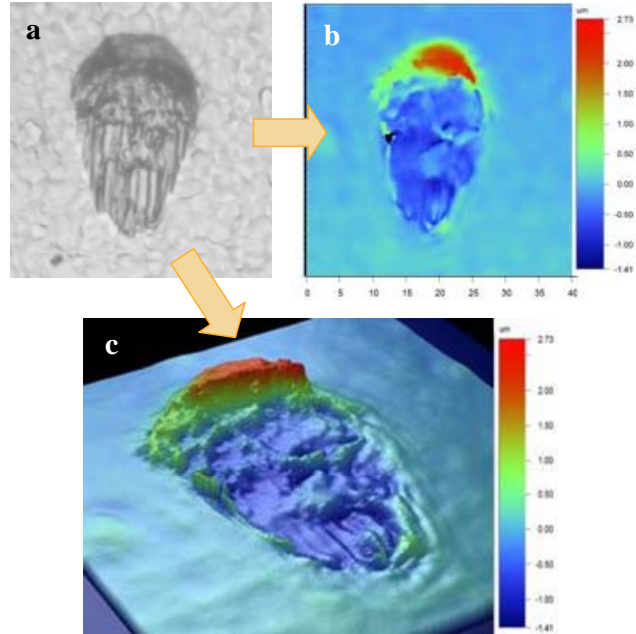


Figure II.2. Views of the aluminium of the pad after EWS: optical microscope top view (a), top view and 3D views of the same mark observed by optical profilometry (b and c).

I.A.2. Expectations and limitations of the technique

The optical profilometry allows a nanometre vertical resolution. However, for optical profilometry, the lateral resolution is limited by the wave length of the light source and the sensors of the camera. For this tool, the scanning velocity and the objectives limits the lateral resolution at 100nm.

The VSI mode allows a 3D mapping of structures whose vertical shift is in the range of micrometres. Since the measurement is done with white light that has a short range of wave length coherence, the fringes of interference are obtained for a very short range of vertical position for each location. As a consequence, steep angles over about 70° cannot be measured.

Finally, optical profilometry is a fast and non destructive imaging technique to observe the surface of the pad and establish mapping profiles. For the current work, it is a very useful apparatus to obtain the profile of a pad damaged after EWS or wire bonding. More specifically, it is further used in the next chapters to map the peeling cases after wire pull tests and to evaluate the scrub mark due to probing and provide insights for the numerical model.

I.B. 3D reconstruction of the pad volume: FIB/SEM reconstruction

In order to obtain the 3D reconstruction of the pad and the whole BEoL volume, an imaging technique that combines the Focused Ion Beam (FIB) and the Scanning Electron Microscopy (SEM) can be employed.

I.B.1. FIB and SEM principles

The Focused Ion Beam (FIB) is used for milling purposes. The overall principle of the technique is similar to the SEM one described hereafter. Ions beam, usually Ga, is used instead of electrons. The interaction between the big and heavy ions and the material mills the latter. This technique is particularly adapted to remove some materials at the microscopic scale.

The scanning electron microscopy (SEM) is based on the sweeping of a surface by an electron beam. The primary electron beam interacts with the atoms of the sample and the electrons loose energy due to repeated random scattering and absorption within a teardrop-shaped volume of the specimen. Hence, according to the type of electron/atom interactions, various electrons are reemitted: secondary electrons, backscattered electrons, Auger electrons and X-rays. For SEM imaging, the low energy secondary electrons are collected to allow the observation of the sample. By using SEM, single or coupled with other techniques, some observation of a surface is obtained.

FIB is mounted in a SEM chamber to combine both techniques. In that case the sample is tilted at about 52° to perform vertical milling and observations simultaneously. The schematic principle of the coupled FIB/SEM tomography is shown in Figure II.3. To obtain the reconstruction of the pad volume, FIB cuts are performed at regular intervals¹, and after each cut, a SEM view is taken. Then, the SEM images can be computed² and the volume is reconstructed in every desired direction.

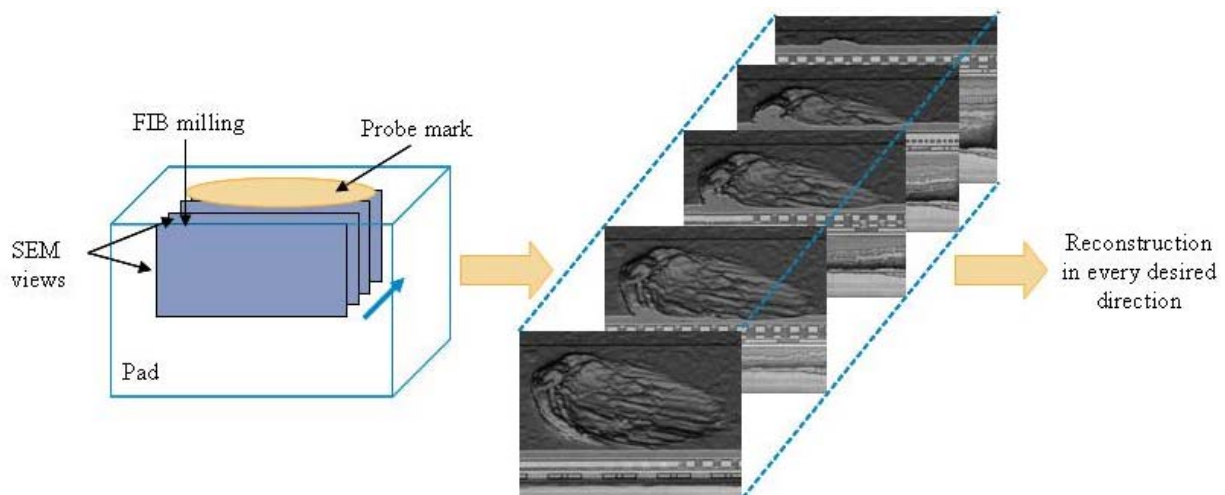


Figure II.3. FIB/SEM reconstruction principle.

I.B.2. Advantages and drawbacks of the technique

The FIB/SEM imaging is particularly efficient and accurate to obtain much information not only on the pad surface, but also to the BEoL structure underneath. Indeed, the observation of the failure under the surface, its path and location are collected. However, this is a destructive and very slow method since only one sample can be treated at a time and both accurate milling and SEM views computing are time consuming. As a consequence, the use of the FIB/SEM technique is limited to research and development and/or for specified investigation purposes.

¹ Every 56nm in the further considered case, but intervals down to 6-10 nanometres can be achieved. Hence, for a $20\mu\text{m}$ wide probe mark, about 350 slices are cuts and views required, which takes several hours.

² In our case, the Amira® software has been used to realign the SEM views.

I.C. Additional tools used to evaluate the pad damages

Additionally to imaging (and particularly to SEM), some dedicated tools can be used to obtain physical-chemical data about the sample. Hence, EBSD and EDX are hereafter presented and their use in the frame of the work underlined.

I.C.1. Electron backscatter diffraction

Electron BackScatter Diffraction (EBSD) is able to characterize and measure grain and sub grain structures for metals, which affect greatly its mechanical properties. The typical main components of an EBSD installation are depicted in Figure II.4. The EBSD apparatus is placed in the SEM chamber. The sample is tilted between 60° and 70° from the horizontal and backscattered electrons from the sample to analyze are collected using detectors providing a diffraction pattern. An image of the diffraction patterns is obtained while many different planes diffract the electrons to form the Kikuchi bands (Figure II.5). Based on the Miller indexes, the bands are indexed and the crystal phases and orientation revealed.

In a publication, Humphreys [Humphreys, 2001] reviewed the most relevant parameters affecting both quality and quantity of the data. These are the material, the pattern acquisition time, the effective spatial resolution and the relative angular precision.

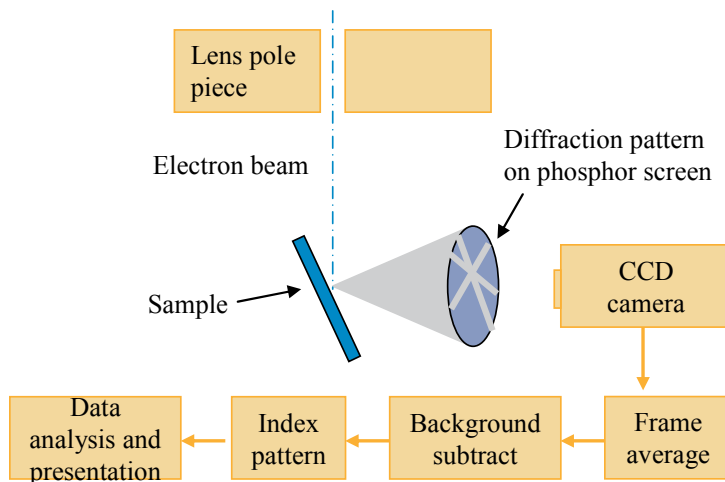


Figure II.4. Schematic diagram of a typical EBSD installation, after [Humphreys, 2001]

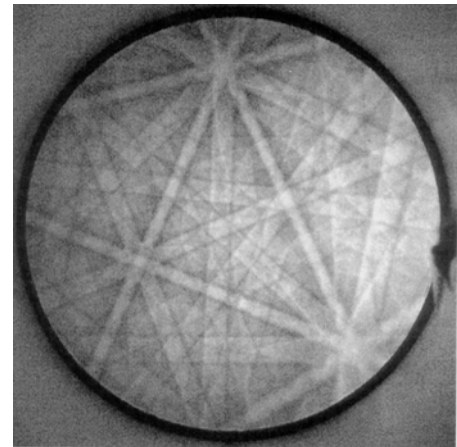


Figure II.5. Example of Kikuchi bands (Nickel). Courtesy of EMSE-CMP.

For this work, this technique is useful to evaluate the modifications of the mechanical properties and the grain sizes and orientation of the various materials implied in the tests and processes. Hence, the damages on aluminium during EWS and the recrystallization of the gold and copper wires for the wire bonding process can be investigated.

I.C.2. Energy dispersive X-Ray

Energy dispersive X-ray spectroscopy (EDX) is used to perform an elemental analysis or a chemical characterization of a sample. The tool is placed in a SEM chamber. The principle is based on a primary source providing an energetic electron that collides with the inner electron of an atom. Hence the electron is removed and replaced by one from an outer shell. Such transition yields a fluorescent photon whose characteristic energy depends on the difference in energy between the initial and the final atomic orbital. The transition from an L shell to a K shell is called $K\alpha$ that is intrinsic to the atoms.

A detector, generally silicon doped with lithium, captures the photons X allowing distinguishing these photons according to their energy. However, some issues related to close transitions energies between the elements can be encountered.

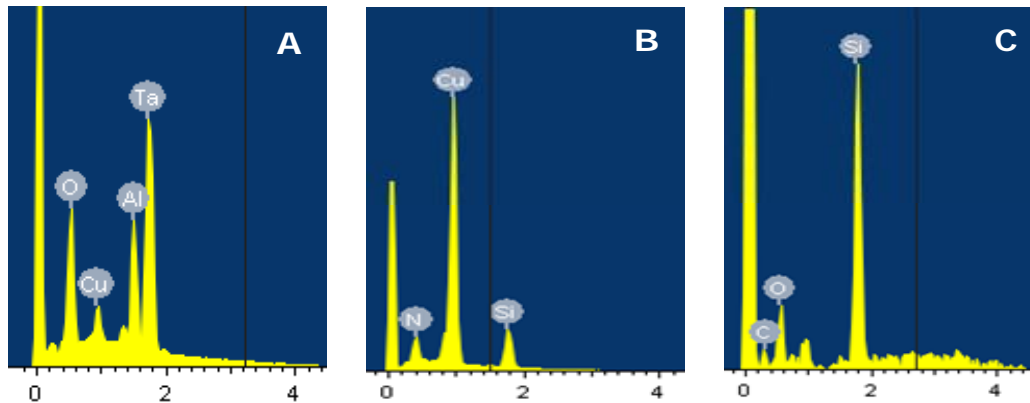


Figure II.6. EDX analyses for surfaces containing O, Cu, Al and Ta elements (a), N, Cu and Si elements (b) and Si, O and C ones (c). Abscissa and ordinate are respectively the energy and the counts.

Some typical results of an EDX analysis are shown in Figure II.6. It enables us to find the chemical composition of a surface, which is, for instance, efficient to find the delaminating interfaces for the peeling cases observed after wire pull tests. The depth of the analysis depends on the input energy of the primary source, but is often limited to few hundreds of nanometres. The peak height is based on the number of counts measured by the detector for a given element.

II. Experimental evaluation of the mechanical robustness of pads: nanoindentation

In the past three decades, a particular focus is paid to the development of new techniques to evaluate the mechanical properties of thin films at the micrometre and nanometre scales. In that frame, the nanoindentation has been developed. This technique derives from the well-known indentation of bulk films. The range of forces applied during depth sensing indentation is between $0.1\mu\text{N}$ and 10N . Such range of forces allows penetrating into the surface to analyze it at very small depth (tens of nanometres) up to tens of micrometers. As a consequence, nanoindentation seems particularly adequate to evaluate the mechanical robustness of pad and reproduce the loading conditions of the EWS and assemblies.

II.A. Principle and apparatus

II.A.1. Presentation of the technique

The general principle is to apply a compressive force (P) or displacement (h) on a sample to extract mechanical properties (Young's modulus, hardness...). The following description concerns standard nanoindentation. Cross section nanoindentation also exists and is used to evaluate adhesion energy between layers, but the technique is not detailed here¹.

The test output is a load-displacement curve that is further post processed to obtain the desired properties.

¹ Measurement of the adhesion energy is performed with cross section nanoindentation. It has been applied to thin films firstly by Sanchez [Sanchez, 1999] and extended to interfacial adhesion energy of pads by S. Gallois-Garreignot [Gallois-Garreignot, 2009].

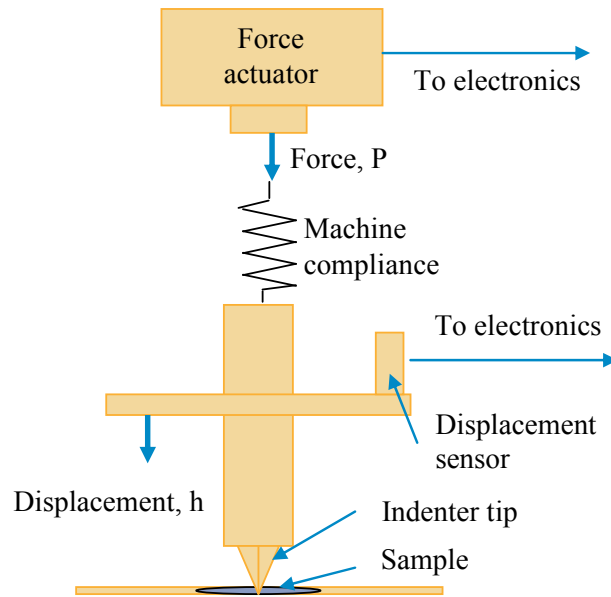


Figure II.7. Schematic representation of the basic components of a nanoindenter.

The schematic representation of the basic components of an instrumented testing system is depicted by the Figure II.7 following Hay et al. [Hay, 2000]. The three major components are the indenter tip with specified geometry, an actuator for applying the force and a sensor for measuring the displacements of the indenter. The low forces can be generated by 3 various means: electromagnetically with a coil and a magnet assembly, electrostatically using a capacitor or with piezoelectric actuators. The displacements are measured by several different ways, whose most common are capacitive sensors and linear variable differential transformers. However, whatever the measurement mean, the displacements include the compliance of the machine. To reduce the effect of the compliance, the machine needs to be calibrated.

II.A.2. Berkovich and cube corner indenter tips

Since nanoindentation involves contact mechanics and elastic-plastic deformation of the tested material, the indenter tip has to be selected carefully. Diamond is often used for these tips since it has high hardness and Young's modulus that minimizes the indenter's compliance. Other materials such as sapphire, tungsten carbide or hardened steel are also used to manufacture the indenter even if they have lower mechanical properties.

Moreover, indenters have various geometries: conical, spherical or pyramidal. In further works, from choice and availability, the Berkovich and cube corner tips are employed to evaluate the structures. Both are three-sided pyramidal tips, as shown in Figure II.8. The geometrical considerations of both tips are displayed in the Table II.1. The Berkovich tip is very commonly used for testing since it has the same depth to area relation than the well-known Vickers four-sided pyramid used for macro indentations. The cube corner tip is very sharp compared to the Berkovich one, and is arranged like the corner of a cube.



Figure II.8. Berkovich (a) and cube corner (b) tips.

	Berkovich tip	Cube corner tip
Centreline-to-face angle, θ	65.30°	35.26°
Area (projected), $A(h)$	$3\sqrt{3} \cdot \tan^2 \theta \cdot h^2 \sim 24.56 \cdot h^2$	$3\sqrt{3} \cdot \tan^2 \theta \cdot h^2 \sim 2.598 \cdot h^2$
Volume-depth relation, $V(h)$	$8.1873 \cdot h^3$	$0.8657 \cdot h^3$
Projected area/face area, A/A_f	0.908	0.5774
Equivalent cone angle, ϕ	70.32°	42.28°

Table II.1. Geometrical considerations for Berkovich and cube corner tips.

The main purpose when using sharper pyramidal indenters (smaller centreline-to-face angle) is to produce larger stresses and strains at the vicinity of the contact. Hence, cube corner tip is often used when investigating fracture toughness as cracking is observed at much smaller loads. The comparison and utilization between the two tips to evaluate pad structures is more widely treated in the following chapters.

II.A.3. General description of the indentation process

A classical load-displacement indentation curve is depicted by Figure II.9. It is obtained when testing silica glass with a Berkovich tip. The curve contains three distinct parts. The first one is the loading stage, up to a maximum force P_{\max} or displacement in the material h_{\max} . It is followed by a holding step, to evaluate the creep of the material. During this plateau, no additional force is applied to the tip that sinks in the material due to its material properties. Then, the unloading stage occurs to remove the indenter. During this step, the material recovers depending on its elastic behaviour and a residual indent is obtained, whose depth h_f can be observed at the end of the test. The Figure II.10 shows a schematic representation of an axisymmetric indentation and the important parameters while performing such testing. The parameter A_c describes the projected contact area between the indenter tip and the sample and h_s is the displacement of the surface on the material during the testing.

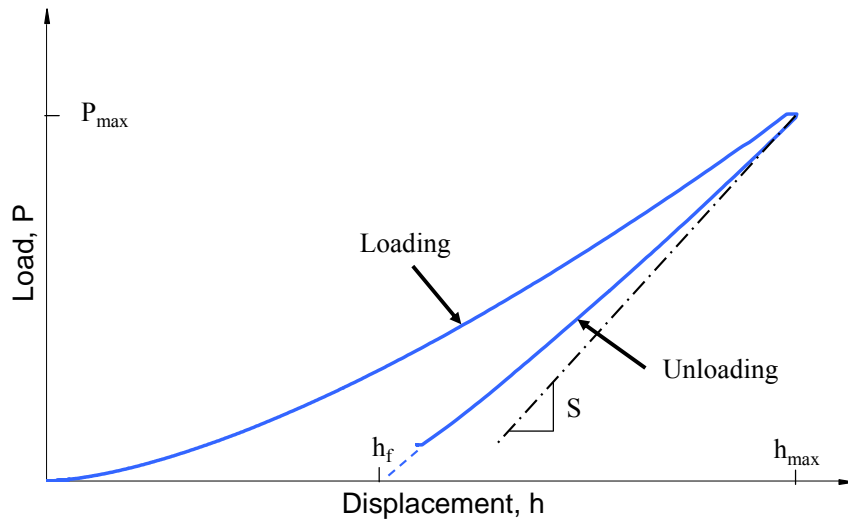


Figure II.9. Indentation load-displacement curve with Berkovich tip in silica glass.

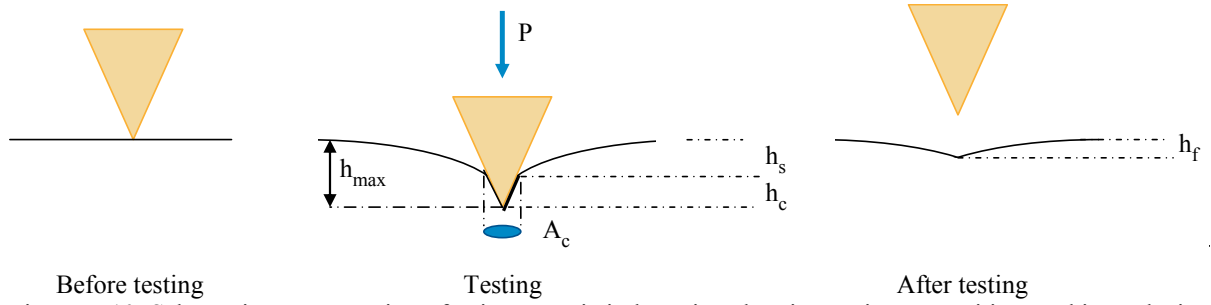


Figure II.10. Schematic representation of axisymmetric indentation showing various quantities used in analysis.

II.B. Determination of material properties

II.B.1. Measurement of hardness and elastic modulus

The hardness is defined as the mean pressure that the material undergoes while loaded. As a consequence, the following equation is obtained:

$$H = \frac{P_{\max}}{A_c} \quad (\text{II. 1})$$

For the nanoindentation technique, Bulychev et al. [Bulychev, 1975] established a relationship to extract the Young's modulus deduced from the slope of the unloading curve:

$$S = \beta \frac{2}{\sqrt{\pi}} E_r \sqrt{A} \quad (\text{II. 2})$$

where β is an indenter geometry parameter, whose value is discussed in section II.B.2.b) and E_r the reduced modulus defined by:

$$\frac{1}{E_r} = \frac{1-\nu^2}{E} + \frac{1-\nu_i^2}{E_i} \quad (\text{II. 3})$$

where E and ν are the Young's modulus and Poisson's ratio of the evaluated material, and E_i and ν_i the same parameters for the indenters. For diamond indenters, the Poisson ratio is 0.07 and the Young's modulus is 1141GPa.

The reduced modulus is introduced in order to take into account the non rigidity of the indenter. Both the hardness and the elastic modulus can be obtained from nanoindentation tests with the help of the Oliver and Pharr method.

II.B.2. Oliver and Pharr model and limitations

II.B.2.a) Description of the model

The first elastic contact models have been developed in the late 19th century by Hertz [Hertz, 1881] and Boussinesq [Boussinesq, 1885]. Hertz analyzed the problem of the elastic contact between two spherical surfaces with different radii and elastic constant. Boussinesq developed a method based on potential theory to evaluate the stresses and displacements in an elastic body that undergoes a rigid and axisymmetric compression under an indenter. Later, Sneddon [Sneddon, 1965] derived relationships between the load and the displacement for flat punch indenter:

$$P = \alpha h^m \quad (\text{II. 4})$$

where α and m are constants. The constant α depends on the material and the exponent m depends on the punch geometry. For flat cylinders, $m=1$, $m=2$ for cones and $m=1.5$ for spheres (considering small displacements) and for paraboloids of revolution. This formula is

valid for elastic materials. Hence, considering the plastic behaviour increases highly the complexity of the problem since additional material parameters are required (yield strength, work hardening...) and constitutive equations are nonlinear. In that frame, Tabor [Tabor, 1948] performed indentation experiments on numerous metals deformed by hardened spherical indenters and finally obtained a well-known approximation between the hardness (H) and the yield stress (σ_y) for metals:

$$H = 3\sigma_y \quad (\text{II. 5})$$

From these considerations, Doerner and Nix [Doerner, 1986] proposed a method to evaluate the slope at the beginning of the unloading step. Their approach is based on the observation that during the initial unloading stages, the contact area remains constant. From this, the methodology to extract the elastic modulus and the hardness of a material is extended by Oliver and Pharr [Oliver, 1992]. It corresponds currently to the most common methodology to obtain these material properties. The Oliver and Pharr method begins by fitting the unloading portion of the load-displacement curve with a power-law relation:

$$P = \alpha(h - h_f)^m \quad (\text{II. 6})$$

where α and m are power-law fitting constants and m is in the range 1.2 to 1.6.

From this, the contact stiffness S is obtained by differentiating the equation (II. 6) and evaluating the result at the maximum depth penetration that gives:

$$S = \left(\frac{dP}{dh} \right)_{h=h_{\max}} = \alpha m (h_{\max} - h_f)^{m-1} \quad (\text{II. 7})$$

The next step is to evaluate the contact depth, which in most cases is different from the maximum displacement into surface due to sink-in or pile-up. The contact area is estimated using:

$$h_c = h_{\max} - \varepsilon \frac{P_{\max}}{S} \quad (\text{II. 8})$$

where ε is a constant that depends on the indenter geometry. For Berkovich and cube corner tips, ε is respectively equal to 0.75 and 0.72.

Finally, the projected contact area is calculated by evaluating an area function at the contact depth h_c . The latter depends on the tip geometry and is given in the Table II.1. However, in practice, indenters have not a perfect sharp shape. To take this into account, the area function of a Berkovich indenter is fitted by the relationship:

$$A_c(h_c) = 24.5h_c^2 + \sum_{i=0}^7 C_i h_c^{1/2^i} \quad (\text{II. 9})$$

where the leading term corresponds to the case of a perfect Berkovich tip and the others describe the derivation from the geometry due to blunting at the tip. For a cube corner indenter, an area function similar to the Berkovich one is obtained, only the leading term is different (given in the Table II.1). This is the main advantage of the Oliver and Pharr method, within the contact area does not need to be measured for each indentation. Hence, prior to indentation tests, the coefficients C_1 to C_8 must be defined through calibration on fused silica, whose Young's modulus is known ($E=72\text{GPa}$).

Once the projected area and contact stiffness are known, the hardness and elastic modulus are obtained from the equations (II. 1) and (II. 2).

II.B.2.b) Limitations of the Oliver and Pharr model

The Oliver and Pharr theory is widely used. However, some assumptions have to be taken into considerations. These assumptions are investigated in a more recent study from the

original authors of the method [Oliver, 2004]. The first one is the elastic properties of the evaluated sample, which leads to sink-in during indentation. However, many materials and particularly metals have elastic-plastic properties and are more prone to form pile-up around the contact. When pile-up is observed, the contact area is greater than that predicted from the method. As a consequence, hardness and elastic modulus can be overestimated by as much as 50% [Bolshakov, 1998]. The Figure II.11 depicts the results from finite elements studies showing the influence of material properties on pile-up. Hence, Bolshakov et al. showed that this phenomenon is important for materials without work-hardening and it is increased when the ratio h_f/h_{max} is over 0.7.

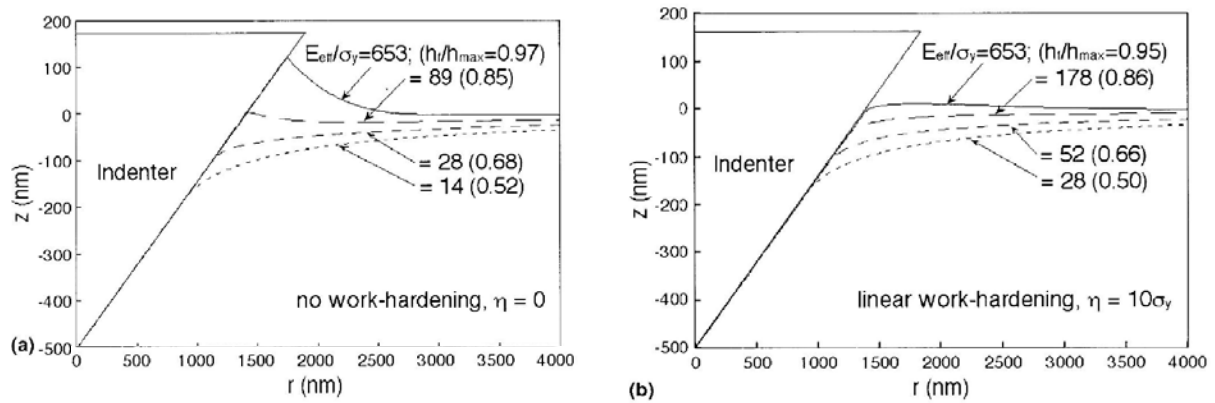


Figure II.11. Finite element simulations of contact profiles under loading with Berkovich tip of non work-hardening material (a) and linear work-hardening material with rate $\eta = 10 \sigma_y$ (b), after [Bolshakov, 1998].

The second source limiting the accuracy of the Oliver and Pharr technique comes from the fitting coefficients in the equations. More precisely, the correction factor β , in equation (II. 2), has a major role on the evaluated hardness and Young's modulus. King [King, 1987] evaluated that $\beta=1.034$ for a triangular indenter. Many recent studies evaluate the value of β . Among others, Hay et al. [Hay, 1999] tried to take into consideration physical and geometrical aspects to establish an equation:

$$\beta = \pi \frac{\frac{\pi}{4} + 0.1548 \cot \phi \frac{1-2\nu}{4(1-\nu)}}{\left(\frac{\pi}{2} - 0.8312 \cot \phi \frac{1-2\nu}{4(1-\nu)} \right)^2} \quad (\text{II. 10})$$

where ϕ is the half included angle (see Table II.1) and ν the Poisson ratio of the indenter, which lead to $\beta=1.067$ for a Berkovich tip. Finally, no clear trend is obtained, but a value close to 1.07 for the latest mentioned indenter seems to provide good results [Oliver, 2004].

II.B.3. Parameters affecting the indentation curves

During nanoindentation on a sample, other sources of inaccuracy need to be taken into consideration. Some materials such as polymers and a few metals are submitted to creep, as described by Chutoba et al. [Chutoba, 2001]. It induces an increase of the displacement at the beginning of the unloading stage. Hence, negative slope can be obtained ($S < 0$) and finally, the material properties are not possible to estimate. In order to reduce the effect of creep on measurement, a holding stage (in further experiments, the holding time is 30 seconds) is applied at the end of the loading as shown in Figure II.9.

Another cause is that indentation tests are time dependent, and strongly influenced by the thermal drift. Thermal variations are related to errors in the measurement of the displacement of the tip. As a consequence, the drift has to be measured and the curve corrected. Usually, the

software connected to the nanoindenter is able to directly correct the load-displacement curves in performing a holding plateau during 15 seconds at the end of the unloading stage.

The strain rate also has a significant effect on the nanoindentation and the extraction of the material properties. In that frame, Wu et al. [Wu, 2006], performed indentations for a large range of strain rates (from 0.0025s^{-1} to 1s^{-1}), and found that this parameter must be kept low in order to avoid having an influence on the measured material properties. Other parameters such as a flat surface are important and more widely described by Hay et al. [Hay, 2000].

II.B.4. Continuous stiffness measurement

The Continuous Stiffness Measurement (CSM) is a technique recently set up to evaluate the stiffness continuously along the loading. This is achieved by applying a small dynamic oscillation for the motion of the indenter, as depicted by Figure II.12. Hence, these oscillations introduce smaller loading-unloading cycles during the test, allowing the determination of the elastic modulus and the hardness at each cycle. The mean to evaluate the stiffness from the measurement is widely described in papers [Li, 2002; Oliver, 2004].

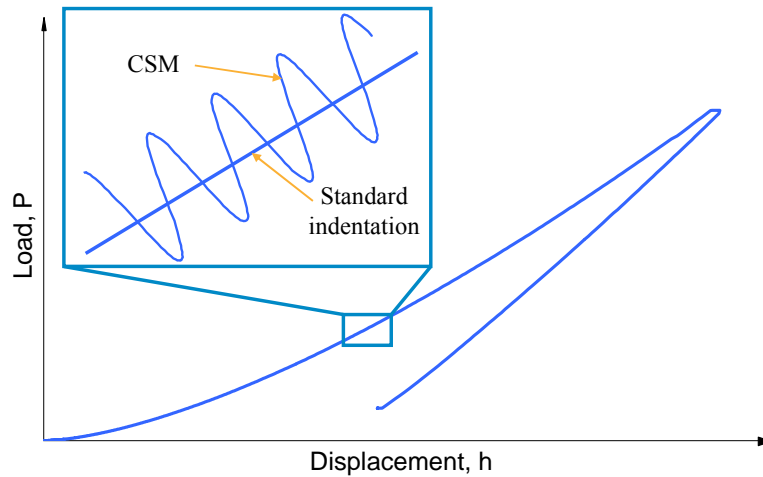


Figure II.12. Typical nanoindentation load-displacement curve with continuous stiffness mode.

The oscillations usually have a frequency of 45Hz and amplitude of 2nm^1 . The effect of these parameters has been evaluated by Wu et al. [Wu, 2006]. For the frequency, they found that higher frequencies reduce the accuracy of the measurements since the loading is no more almost static. Moreover, the amplitude is found to have little influence on the measurements.

II.C. Considerations for the indentation of multilayer structures

The technique and the tips used during the nanoindentation tests have been presented previously. The uncertainties during the testing due to pile-up, creep, thermal drift, etc have been shown to be considered with great care to obtain reliable results. The Berkovich tip is commonly used in nanoindentation, but, based on geometrical considerations; the Cube Corner (CC) tip is sharper than the Berkovich one. Thus, for a given depth, the indented volume is smaller, and implies larger strains.

For application to the microelectronic, the nanoindentation technique is performed on multilayer structures. As aforementioned, they are made from an alternation between metal and oxide films, which have very different mechanical properties and material behaviours.

In the literature, several studies evaluate the effect of the substrate (bi-layer indentation) on the material properties of the thin film. Usually, to obtain E and H of the film independently

¹ These values are Agilent/MTS recommendations for the CSM mode with the XP nanoindenter.

on the substrate, the indentation depth is less than 10% of the film thickness, as described by Hay et al. [Hay, 2000] or by Fischer-Cripps [Fischer-Cripps, 2006]. Moreover, the former showed that in some cases, depending on the elastic and plastic behaviours of the layers, values for E and H can be trusted until 25% of the film's thickness. The latter describes the common errors while interpreting the nanoindentation curves and extracting the material properties. He underlines that piling-up induces very high uncertainty in the results to extract E and H , especially for films on substrates. Hence, it is recommended to use E and H as qualitative quantities for comparison purposes.

In our case, the aluminium film is on a harder oxide substrate. In that frame, Soare and co-workers [Soare, 2004] evaluated through experiments and simulation the effects of creep and pile-up for aluminium on silicon and bulk aluminium. They found that in both case, pile-up, even using a Berkovich tip is significant, modifying the determined Young's modulus and hardness. The hardness of aluminium on hard substrate is higher than on bulk aluminium due to the substrate effect, even at small depth. Chen et al. [Chen, 2005] performed nanoindentation using the Berkovich tip on glass and silicon substrates. They also observed the effect of the substrate on the measurement. In their study, Saha and Nix [Saha, 2002] showed the substrate effect while considering aluminium thin films. The Figure II.13 and the Figure II.14 show respectively the load-displacement curves and the measured hardness, according to the Oliver and Pharr technique. The effect of a harder substrate is clearly highlighted. They found that the increase of the aluminium hardness on a harder substrate is not very important because the plastic deformation is contained within the film and the substrate yield plastically only when the tip penetrates in.

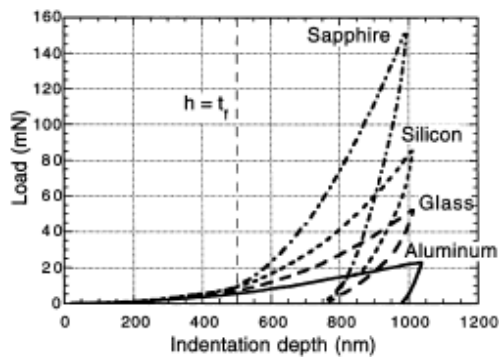


Figure II.13. Load vs. indentation depth for $0.5\mu\text{m}$ films on sapphire, silicon, glass and aluminium substrates, after [Saha, 2002].

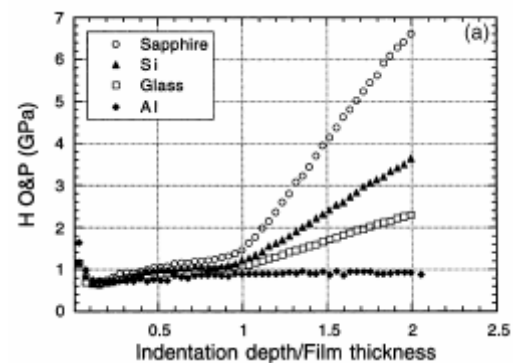


Figure II.14. Hardness for the $0.5\mu\text{m}$ Al film on the substrates as a function of the depth normalized by the film thickness, after [Saha, 2002].

Another target of the nanoindentation on pads is to induce cracks in the oxide layer below the aluminium. Due to the sharpness of the cube corner tip, the stress induced is high, and can lead to cracks in brittle films. In the literature, several studies have been reported to evaluate low- k oxides. Morris et al. [Morris, 2008] found that in some cases the Berkovich tip is not acute enough to generate radial fracture in low- k films. On the contrary, the use of the cube corner tip allows the cracks generation. Similarly, Nay et al. [Nay, 2004] and Fujikane et al. [Fujikane, 2008] demonstrated that the cube corner tip induces more likely cracks in the low- k oxide thin film even for very low load and displacement into the surface. The effect of the geometry of the indenter tip is considered by Jang et al. [Jang, 2008] who demonstrated that reducing the equivalent cone angle of the indenter induces an increase of the displacement into surface in Si and Ge films for a given load. Additionally, the ability of the tip to induce cracks in a brittle film increases as well.

The cracks observed during the nanoindentation with the CC tip are characterized by discontinuities in the load-displacement curve, as obtained by Simonyi and co workers

[Simonyi, 2005]. Similarly, Chen et al. [Chen, 2006] showed that nanoindentation under displacement control is suitable to investigate small cracks in submicron films. Furthermore, the discontinuity caused by the failures is correlated to an energy dissipated by the film.

The ability of nanoindentation to compare various oxides has been studied by Vella et al. [Vella, 2003], who evaluated the mechanical properties and the fracture toughness of low-k films (OSG), and shown that nanoindentation is an efficient tool to compare various oxide films. For the mechanical properties, they observed an important substrate effect. Yet, recent studies [Hay, 2009] have shown the possibility to have substrate independent modulus even for low-k film as thin as 200nm in combining simulation and experiments.

Finally, it seems that the nanoindentation raises several issues for the measurement of mechanical properties on multi-layers stacks. Additionally, in our case, the aluminium, which is prone to very important pile-up, and the objective to obtain cracks in the oxide thin layer, harden the pad study. At the same time it reduces the reliability of the hardness and the Young's modulus measurements due to the multilayer system.

II.D. Nanoindentation experiments for pad robustness evaluation

In the literature (later described in chapter 3 to 5), many papers deal with nanoindentation studies and its use as mean to evaluate the material properties of thin films. However, several issues are still faced such as the effect of the substrate on the measurement. The use of nanoindentation to reproduce the EWS test and wire bonding assembly to compare the pads robustness in which several thin films stacked is still limited. Nanoindentation corresponds to a simplified test with respect to EWS or WB because only a compressive load is applied to the pad. In further work, the ability of nanoindentation to reproduce the failures in the oxide layers and to evaluate the pad robustness is questioned. Additionally, material properties of the pad are also measured to provide a background for modeling.

III. EWS numerical model

In the frame of reproducing the probing tests, a dedicated Finite Element (FE) model is developed to mimic the probing. The model is developed using the explicit solver¹ ANSYS LS Dyna release 11.

The reproduction of the actual EWS is very complex and related to many challenges. As previously mentioned, during the test, a contact is established between the hard probe and the aluminium of the pad. The contact and its related non linearity have to be modelled and also the high damages undergone by the aluminium and oxide layers. Moreover, probing is a transient phenomena and it implies dynamics of the needle during and after the contact. The use of an explicit solver perfectly matches the expectations since it has a higher ability to manage the contacts and take into account the dynamical effects compared to an implicit solver². The general equation of the motion (u: displacement) of the needle, solved by the explicit solver is defined by:

$$[M]\{\ddot{u}\} + [C]\{\dot{u}\} + [K]\{u\} = \{F(t)\} \quad (\text{II. 11})$$

where [M], [C] and [K] are respectively the mass, damping and stiffness matrixes.

In order to evaluate the robustness of the pads, a multi-scale issue is raised considering the difference between the size of the probe (in the range of the millimetre) and the size of the pad and the detailed in the stack below (in the range of the micrometre and even below

¹ See chapter 5 for comparison between explicit and implicit solving methods.

² The probing model for an implicit solver scheme is described and discussed in chapter 5.

depending on the details). Additionally, the analysis requires a three dimensional model to investigate the effect of pad layouts.

III.A. Model description

III.A.1. Geometrical model

Similarly to EWS, the geometric FE model is made from two parts, which are the probe and the pad. Only half of the model is simulated due to symmetry plane. The Figure II. shows these parts and highlights the various scales faced during the probing modeling. Both parts are meshed with SOLID164 quadrangle elements. These are 8 nodes 3D elements with one integration point (reduced integration option).

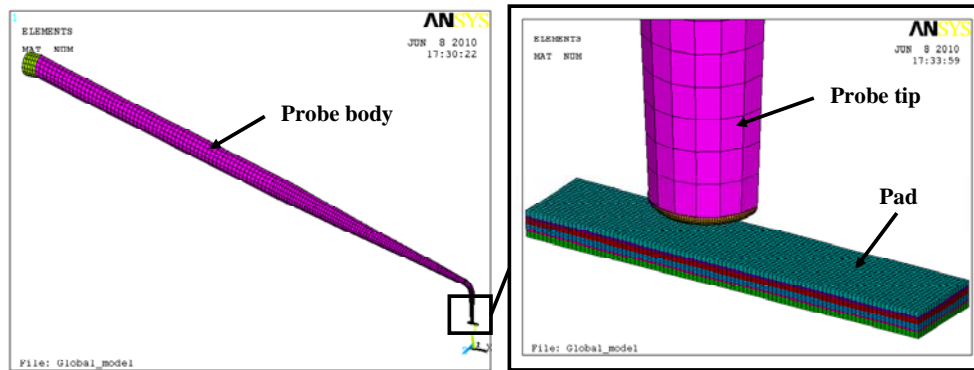


Figure II.15. Model for level 2 probe geometry (left) and zoom at the probe's tip and the pad (right).

III.A.1.a) Probe side of the model

For the probes, four various geometries are considered as shown in Figure II.16. They correspond to existing needle levels as part of the cantilever probe cards H9 and Helium+ (respectively for the testing on CMOS 120nm and CMOS 45nm technology nodes), manufactured by Technoprobe, one of the STMicroelectronics provider. The geometries of the needles for the two probe cards slightly differ. Thus, the geometries and material properties of the probes are given by the supplier.

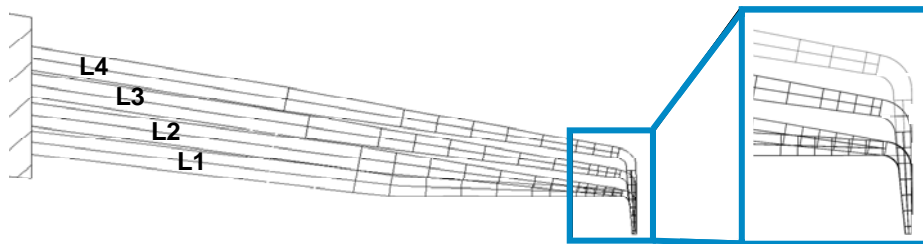


Figure II.16. The four various geometries of probe and probe tip studied for the H9 probe card.

The finite element model of the needle side has been previously developed and described by Cacho et al. [Cacho, 2007]. Additionally to the probe, to get closer to the EWS test, an effective volume is added to take into account the rigidity of the probe card at the probe's end. This volume is required because the needle is fixed using an epoxy resin that has a low rigidity.

III.A.1.b) Pad side of the model: CMOS 120nm pad

The typical components and manufacturing of the pad have been previously described in the first chapter and are shown in Figure II.17 (a). A side view of the numerical model corresponding to this pad is depicted in Figure II.17 (b). Two main differences exist between the model and the actual pad. Firstly, the barrier layers such as TaN/Ta are not represented in the model. Indeed, it is assumed that this very thin layer (30 to 40nm) has little effect on the overall robustness of the structure. Moreover, the thickness of this layer is constant whatever the pad stacking and design options. Indeed, if the pad open oxide and the tantalum layers are homogenized, the result is very close to the material properties of the oxide alone, and further experimental work confirms that both layers have a brittle behaviour.

Secondly, some layers or details are not represented. The top copper layer (M4) is described by the corresponding homogenized properties¹. The third level of metal (M3) and the oxide layers right above (IMD3) are represented as one layer. The stack made from all the oxide and metal layers bellow IMD2 (M1, M2...) are also regrouped in one layer. Here, it is assumed that most of the mechanical issues during EWS are located on the top of the pad.

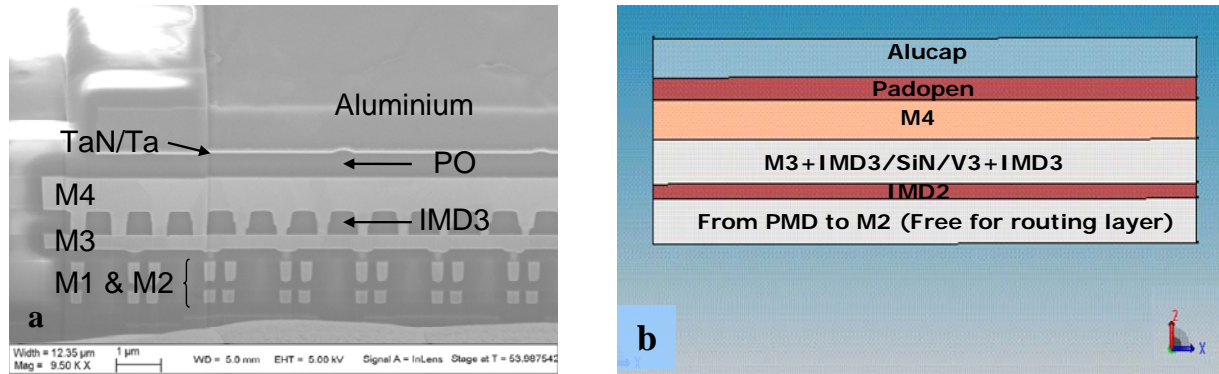


Figure II.17. SEM cross section view of a CMOS 120nm pad (a) and side view of the corresponding model (b).

III.A.2. Material Properties

III.A.2.a) Homogenization technique

In the pad, the various designs necessary to perform the electrical connection can be smaller than 1 micrometre. Hence, it is not possible to insert them directly in the model and the use of homogenized properties is required to take into account the mechanical properties related to these small patterns. The numerical method described had been developed specially for the pad evaluation by Fiori et al. [Fiori, 2005]. Due to the patterning of the pad, isotropic elastic material properties are not relevant in most of the cases and anisotropic equivalent mechanical parameters are needed. The pad structure is often periodic, thus a representative unit cell can be extracted. With the help of boundary conditions with respect to the model, 6 loading cases are applied to evaluate the 9 equivalent orthotropic properties from the Hooke's law [Ansys LSDyna, rel. 11]:

$$\{\varepsilon\} = \{\varepsilon^{th}\} + [D]^{-1} \{\sigma\} \quad (II. 12)$$

where $\{\varepsilon\}$ is the total strain vector, $\{\varepsilon^{th}\}$ the thermal strain vector that is null in the present case, $[D]$ is the elastic stiffness matrix and $\{\sigma\}$ the stress vector.

Thus, by applying successively tensile loads along each axis, the Young's moduli and Poisson's ratios are deduced. The same operation with shear loads gives the shear moduli G .

¹ The several designs for the top copper layer investigated during this work are later described in the chapter 4.

The homogenization method can be used for one or several layers. The results obtained by Fiori et al. applied to wire bonding show that both can be used with good results, and the former is more accurate but increases the CPU consumption.

III.A.2.b) Material properties

The probe's needle is mainly made from tungsten and its material properties are employed, though it is actually made from a WRe alloy. Bulk tungsten has a Young's modulus, a Poisson's ratio and a density of respectively of 390GPa, 0.29 and 19.3g/cm³. The use of elastic material properties for the probe is based on its mechanical properties compared to those of aluminium and the fact that it is almost not damaged during a test. This assumption is validated by the fact that several thousands of touchdowns are made with these needles as described in the section II.A.4 of the first chapter. Moreover, Liu et al. investigated by means of a numerical model the stress in the probe during one test and found it below the yield stress of the tungsten (1.899GPa) [Liu, 2006].

The aluminium and oxide layers are provided with isotropic elastic material properties. The use of plastic material properties for aluminium increases highly the CPU consumption and can lead to convergence issues, so it is not considered in this part. The density is inputted into the model since the explicit solver is able to reproduce inertial effects. For layers containing copper and oxide, the homogenization technique is applied to get orthotropic material properties. For these close to the surface of the pad, the method is applied to a single layer, and for the deeper ones, the homogenization is applied to a stack of several layers. All the material properties used for the model are shown in Table II. 2. The orthotropic properties in the x-y plane of the pad are different due to the small patterns in the various layers, such as vias.

Material	Young's modulus (GPa)	Poisson's ratio	Shear modulus (GPa)	Density (g/cm ³)
Aluminium	69	0.34	/	2.7
SiO ₂ (USG) (pad open and IMD2)	63	0.23	/	2.2
Top copper layer (M4) – Slots ¹	E _x =120 E _y =115 E _z =120	v=0.25	G _{xy} =45 G _{xz} =43 G _{yz} =45	7.27
Top copper layer (M4) – Lines ¹	E _x =132 E _y =132 E _z =125	v _{xy} =v _{xz} =0.26 v _{yz} =0.25	G _{xy} =52 G _{xz} =49.3 G _{yz} =49.3	8.96
M3 / IMD3	E _x =81.1 E _y =78.8 E _z =74.7	v=0.23	G _{xy} =49.3 G _{xz} =52 G _{yz} =49.3	3.68
Free for routing layers	E _x =98.3 E _y =98.8 E _z =94.4	v=0.28	G _{xy} =37.7 G _{yz} =35.9 G _{xz} =35.8	4.96

Table II. 2. Elastic isotropic and orthotropic material properties used for the pad.

III.A.3. Loading conditions and contact checking

During EWS tests, the chuck is raised to establish the contact with the needles of the probe card. Hence, for the model, in a first attempt (i.e. during the calibration step described in the section III.B.1.b)), the loading is applied to the pad. The velocity of the calibration

¹ The M4 layouts are described and evaluated in the section C of the fourth chapter.

experiments, whose profile during one pass is shown in Figure II.19, is used to mimic the probing. Such a velocity profile provides a similar displacement (or overdrive) of the pad to the one observed during the electrical tests. The impact velocity is fixed at $827.5\mu\text{m/ms}$.

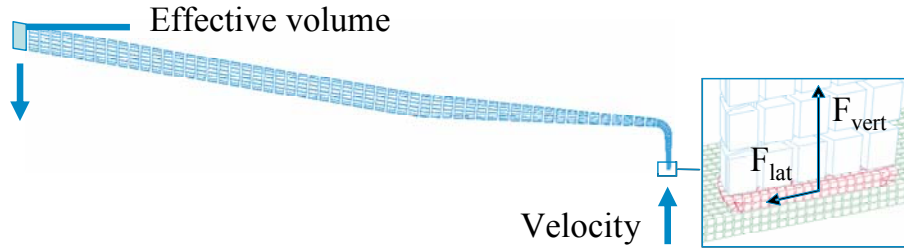


Figure II.18. Input and outputs for simulations.

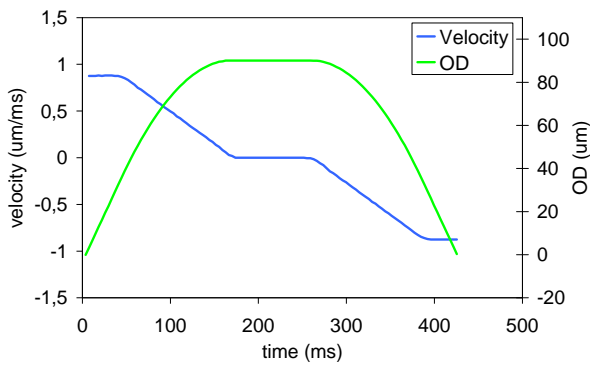


Figure II.19. Velocity and overdrive as a function of time.

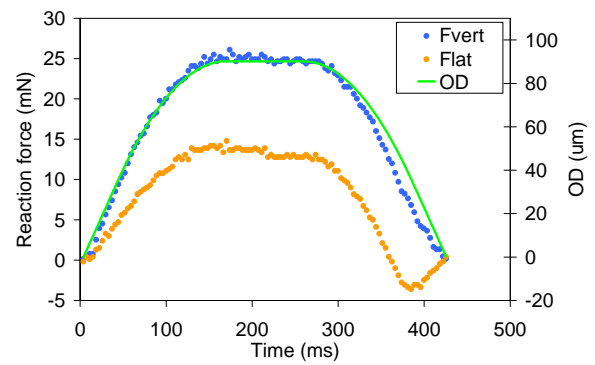


Figure II.20. Vertical and lateral reaction forces and overdrive as a function of time.

Hence, the first control after the solving the model is to check that the OD applied is the one desired. The outputs of the simulation to verify that the contact is well managed are the vertical and lateral reaction forces between the needle and the pad. The Figure II.20 shows these forces during one touchdown. Thus, during the loading, holding and unloading stages, the vertical reaction force is proportional to the imposed displacement and remains constant when the maximum overdrive is reached. The variations of the lateral reaction force are more complicated. During the loading stage, the force is proportional to the sliding distance due to the elastic deformation of the probe. During the unloading step, the force direction is inverted. At the beginning of the unloading, the sliding distance is null. It is due to the elastic recovery of the probe body, slightly bended during the loading step. After that, the friction force is opposite to the sliding motion and finally the lateral force decreases to null when the contact is lost.

In order to simplify the model and avoid issues during pad evaluation studies, the velocity is applied to the effective volume. Velocities applied have opposed signs and it is checked that they lead exactly to the same reaction forces undergone by the surface of the pad.

III.A.4. Physical and numerical parameters

A couple of physical and numerical parameters can have an influence on the results given by simulations. They are the effects of the mass scaling, the damping of both the probe and the contact, the friction coefficient, the meshing, the reduced integration and the effective volume. Using explicit solver for simulations is highly CPU time consuming. Hence, a good trade-off has to be obtained between resource consumption and solution accuracy.

III.A.4.a) Effect of the meshing

The meshing size has an influence because increasing the number of finite elements implies a simulation model closer to the physical model. But the CPU time increases exponentially as well. Hence, a good trade-off between accuracy of the simulation and CPU time consumption is required. The convergence of the model has also to be checked to simulate accurately both vertical and lateral forces. The simulations of the influence of the number of elements were performed without effective volume. The level 1 probe was used for this set of experiments. The overdrive is fixed at 100 μ m.

Elements number	Normalized vertical force (%)	Normalized lateral force (%)
230	63,2	62,9
562	71,3	71,1
807	81,0	80,7
1691	91,9	91,9
2435	96,4	95,9
6645	100,4	100
8894	100	100

Table II.3. Normalized lateral and vertical forces vs. number of finite elements in the probe.

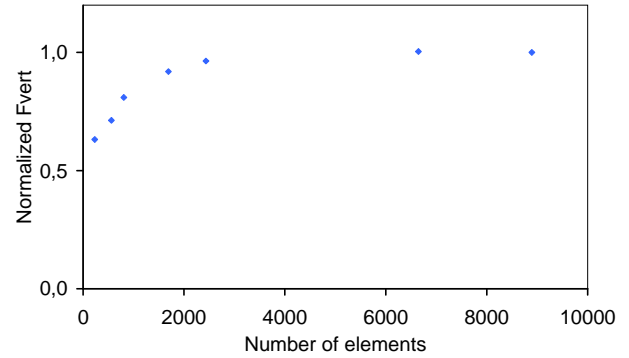


Figure II.21. Normalized vertical force vs. number of elements.

The Figure II.21 highlights that there is a convergence of the model by increasing the number of finite elements (and as a consequence decreasing the size of each element). As shown in the Table II.3, over 6000 elements, a constant value of the vertical force is reached (and similarly the lateral force for reasons explained hereafter). Nevertheless, this number of finite elements is highly CPU time consuming and thus for further simulations, the model with 2435 elements is chosen. This number of elements corresponds to about 96% of the maximal vertical and lateral forces and is far less CPU time consuming.

III.A.4.b) Effect of the mass scaling

The mass scaling effect is a simulation parameter that has a huge influence on the CPU time. The use of mass scaling allows a reduction of the CPU time consumption by increasing the minimum time step. The time step compels the explicit algorithm. This minimum time step is based on its material properties and characteristic length:

$$\Delta t_{\min} = \frac{l_{\min}}{c} \quad \text{and} \quad c = \sqrt{\frac{E}{(1-\nu^2) \cdot \rho}} \quad (\text{II. 13})$$

where l_{\min} is the length of the smallest element, E , ν , ρ are respectively the Young modulus, the Poisson ratio and the density of the materials.

The previous equation can be rearranged to find the required density of each element for a desired time step size:

$$\rho_i = \frac{(\Delta t_{\text{specified}})^2 \cdot E}{l_i^2 \cdot (1-\nu^2)} \quad \text{for the element } i. \quad (\text{II. 14})$$

Mass scaling consists in adding the corresponding mass to these elements, which reduces the solving time. But it can also invalidate the results in case of abuse. Thus, a Differential

Time step Mass Scaling (dtms) has to be fixed for simulations. A high time step parameter ($dtms = 1.10^{-3}$) gives a very low CPU time consumption by reducing the number of iteration cycles but it also reduces the accuracy of the results. On the contrary, a low time step parameter ($dtms = 1.10^{-5}$) gives a very precise result but the solving requires a huge CPU time consumption.

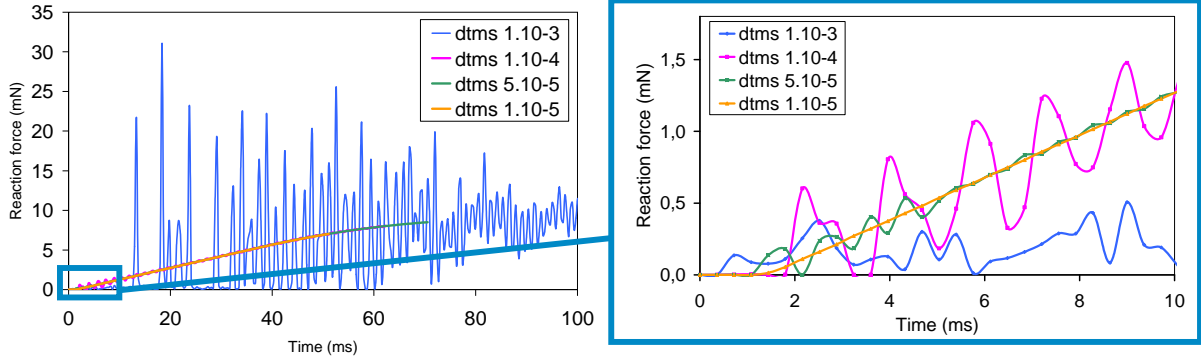


Figure II.22. Effect of mass scaling on the vertical reaction force during the loading stage (left) and for the first ten milliseconds after the contact (right).

The Figure II.22 shows the mass scaling effect on the management of the contact during the simulation. For a desired minimum time step of 1.10^{-5} (orange curve), the curve of the vertical reaction force simulated is smooth even just after the contact and shows no resonance effect of the probe. This time step is good with respect to the probing test but is very high CPU time consuming. On the contrary, a time step of 1.10^{-3} (blue curve) gives a wrong result and finally doesn't simulate accurately the interactions between the probe tip and the pad. Time steps of 1.10^{-4} (violet curve) and 5.10^{-5} (green curve) give intermediate results and are quite good trades-off between CPU time consumption and accuracy of the result. Those 2 mass scales are further considered to evaluate the other parameters. The former is chosen for results that require medium accuracy (typically for the determination of the Young's modulus of the effective volume) and the later is used when quite high precision of the results is necessary (meshing study).

III.A.4.c) Effect of the friction coefficient

The friction coefficient (noted μ) is the force resisting to the lateral motion of the probe, and is defined by the Coulomb equation:

$$F_{vertical} = \mu * F_{lateral} \quad (II. 15)$$

In the model, it is the single parameter that can be fixed to obtain the lateral reaction force and describe the lateral displacement of the needle. The value of the friction coefficient is determined from the calibration performed by the supplier (described in III.B). For the model of the H9 probe card, the friction coefficient was fixed at 0.096. The one of the Helium+ probe card is much higher (about 0.58).

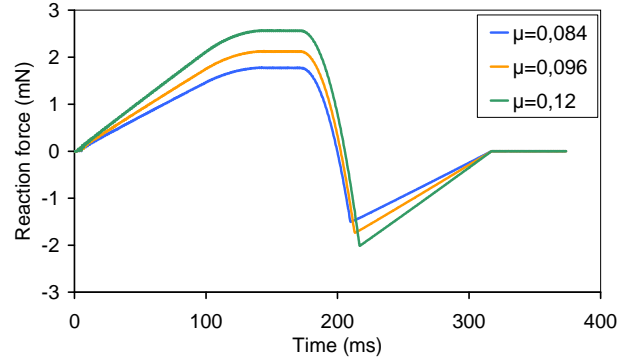


Figure II.23. Evolution of the lateral reaction force versus the friction coefficient for experimental measurements.

The experimental uncertainty of this coefficient sometimes reached 20%. Hence, for the simulations, friction coefficients equals to 0.084, 0.096 and 0.12 are compared for the calibration of the H9 probe card. The simulations are performed for an overdrive of 100 μ m and level 2 needle geometry. Therefore, the vertical reaction force is the same for each of the simulations. The lateral reaction forces increases linearly with the increase of the friction coefficient, as depicted in the Figure II.23.

III.A.4.d) Effect of the effective volume

During the first calibration attempts, both the lateral and vertical reaction forces reached are very high compared to those given by experiments from Technoprobe. Hence, it appears necessary to add an effective volume to better fit experiments and simulations. The forces measured cannot be totally attributed to the probe needle and the assumption that the blocking system at the end of the needle is totally rigid (i.e. infinitely stiff) is not valid because a small quantity of the force is absorbed by the probe card. The probe end is linked to the probe card by some glue and an epoxy resin whose mechanical behaviours are unknown. Hence, an effective volume is required at the probe end. The volume of the added part is fixed at $V_{eff} = 0,00188\text{mm}^3$ similarly to Cacho et al [Cacho, 2007]. An increase of the volume is related to a decrease of the reaction forces. To fit the experimental results, the Young's modulus of the effective volume is evaluated.

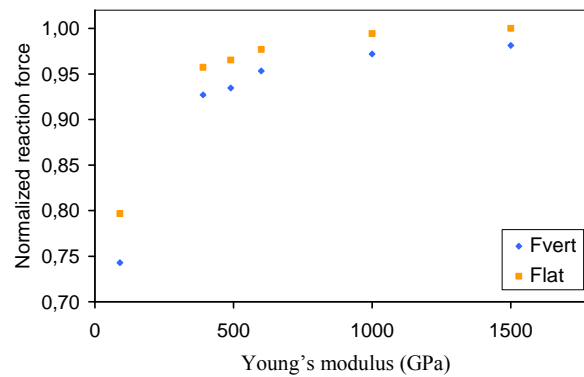


Figure II.24. Evolution of the lateral and vertical forces versus Young's modulus.

The evolution of both the lateral and vertical forces with respect to the elastic modulus is evaluated as shown in Figure II.24. The highest value of the reaction forces corresponds to the case without effective volume (i.e. the case of an effective volume with infinite elastic modulus). An increase of the Young's modulus raises both reaction forces.

The elastic modulus of the effective volume is deduced from the curves giving the lateral and vertical forces versus the Young's modulus, for the various probe geometries. It is assumed that all the levels require the same Young's modulus for the effective volume. This assumption is checked later when the fitting between simulations and experiments is performed. Hence, for those geometries, the study was performed for the 3 different overdrives (50 μ m, 75 μ m and 100 μ m). As a consequence, the elastic modulus is fixed at 90GPa.

III.A.4.e) Effect of the reduced integration

The elements with reduced integration use one integration point, located in their centre (for SOLID164 elements). On the contrary, fully integrated elements are commonly used for implicit modeling, and usually have 8 integration points. The main advantage of using reduced integration is the CPU time saving by minimizing element processing. In addition, single integration elements are also extremely robust in case of large deformations. However, reduced integration elements are prone to have hourglassing. It is a zero-energy mode of deformation that oscillates at a frequency much higher than the structure global response, as shown in Figure II.25. Hourglassing modes result in stable mathematical states that are physically not possible. They typically have no stiffness and give zigzag deformation appearance to a mesh as depicted by Figure II.26. The occurrence of hourglass deformations in an analysis can invalidate results and have to be minimized or eliminated. The overall hourglass energy has to stay below 10% of the internal energy of a model, otherwise there is likely an issue.

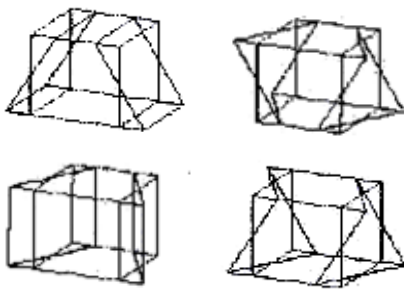


Figure II.25. Zero energy deformation for reduced integration elements, after [Ansys LSDyna, rel. 11].

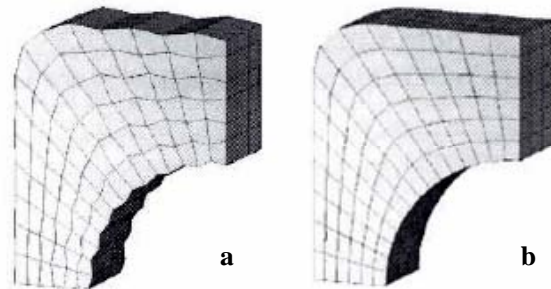


Figure II.26. Model with (a) and without (b) hourglass after [Ansys LSDyna, rel. 11].

Hence in the current model, the reduced integration is selected, this assumption has to be checked after each simulation, and the validity for this model is discussed in section III.B.1.c).

III.A.4.f) Effect of the damping

Two kinds of damping exist: the damping due to the contact that is a numerical parameter and the damping due to the materials that is rather a physical parameter. In the current study, both damping types have limited effect on the final result of the simulation because the convergence of the reaction forces is obtained (see paragraph III.A.4.a)). By modifying the damping values, there is no modification of both vertical and lateral reaction forces.

III.B. EWS numerical model set up: calibration

The main interest of the study is to fit experiments and simulations in order to obtain a reliable numerical model. Assumptions (mesh size, Young's modulus of the effective volume ...) have been performed and are now faced to experiments.

III.B.1. Experiment and calibration for CMOS 120nm pad model

The finite element model for the CMOS 120nm technology node is calibrated with H9 cantilever probe card, whose needle geometry has been previously shown.

III.B.1.a) Experimental measurements of the reaction forces

A specific tool is developed by Technoprobe since it is not possible to perform measurements in-situ. The apparatus is shown in Figure II.27. To measure the vertical and lateral reaction forces, a vertical displacement is applied to one needle. Hence, this procedure allows obtaining the forces for the four needle geometries and for various overdrives and number of touchdowns.

From these experimental measurements and the Coulomb's law of friction, the friction coefficient is estimated being equal to 0.096.

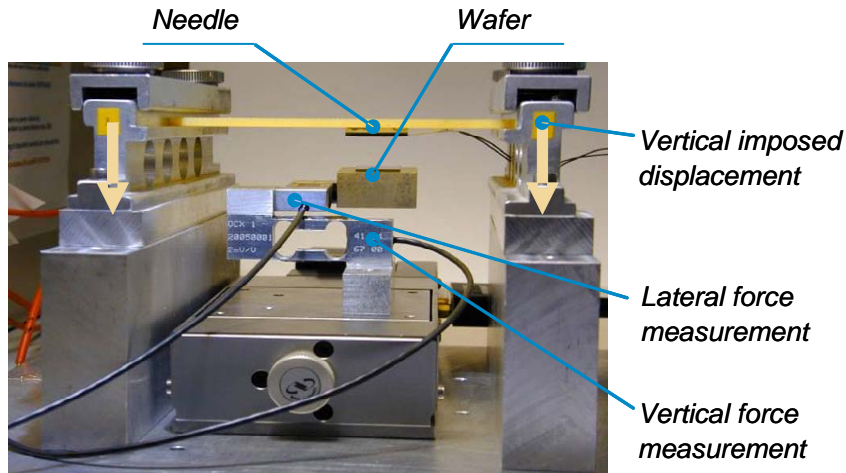


Figure II.27. Experimental tool set up by Technoprobe to measure the vertical and lateral reaction forces.

These measurements allow a fair reproduction of the actual process as depicted by the Figure II.28. To compare the results from the numerical model and the experimental calibration, the maximum reaction forces, observed during the holding stage are extracted for each parameter. All the values are gathered in Figure II.29 after one touchdown. Hence, it is shown that the vertical reaction force increases linearly with the increase of the overdrive. The needle geometry has a slight effect on the force and a ranking of the needles according to the force they induce to the pad is established:

$$level\ 1 < level\ 2 < level\ 4 < level\ 3 \quad (II. 16)$$

For the lateral forces, the effect of the overdrive is the same. However, some experimental maxima such as the overdrive 50 μ m for the needle level 1 are questioned, since the measurement is different from the linear increase expected. Indeed, the lateral force values are of the order of the millinewton that is the order of magnitude of the experimental resolution and is probably the reason for the unexpected values obtained.

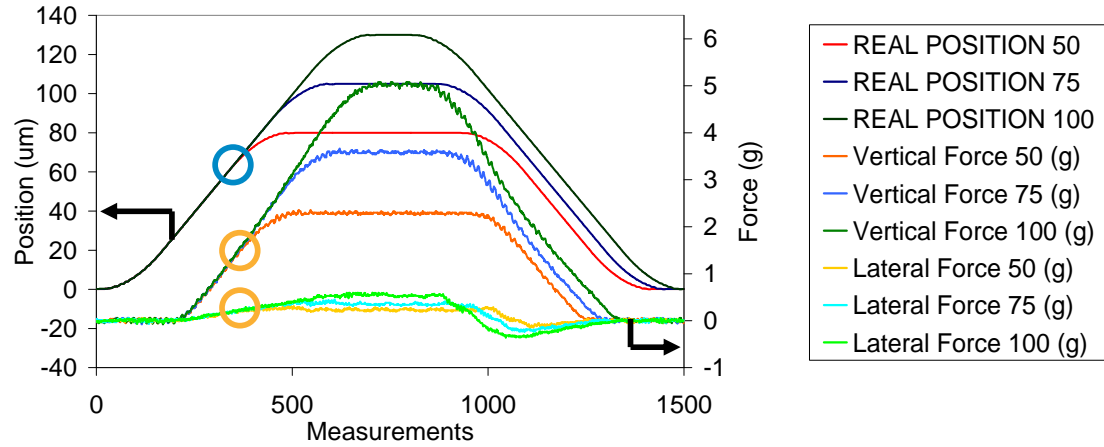


Figure II.28. Input (position, blue circled curves) and outputs (forces, orange circled curves) from Technoprobe.

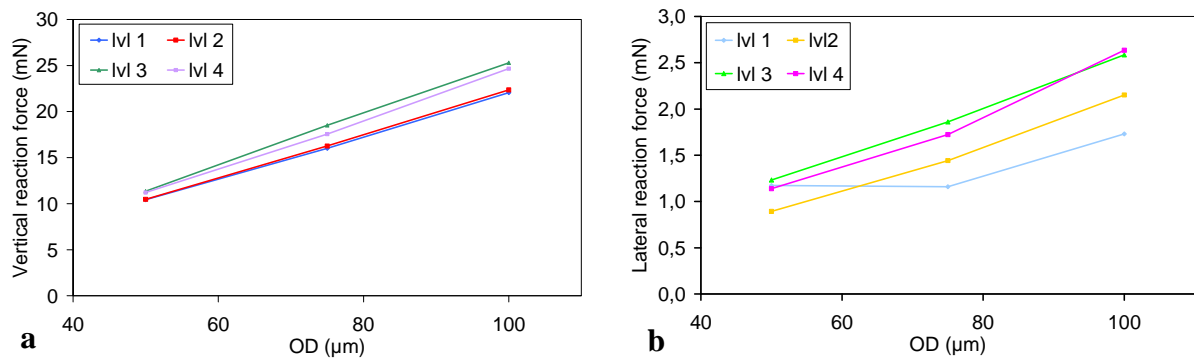


Figure II.29. Evolution of the maximum vertical (a) and lateral (b) reaction forces measured experimentally for various overdrives and for the four needle geometries (levels one to four).

III.B.1.b) Calibration of the numerical model

The same parameters than the one evaluated during the experiments are simulated with the previously described finite element model. Hence, the 4 geometries of probe are evaluated for 3 different overdrives (50μm, 75μm and 100μm) with an effective volume at the end of the probe needle as defined before (elastic modulus of 90GPa). The Figure II.30 shows that the numerical model is able to reproduce the behaviour of the whole experimental curve for both vertical and lateral reaction forces.

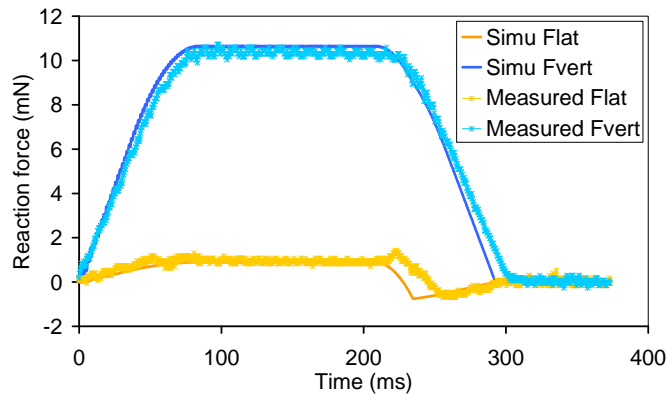


Figure II.30. Simulated and measured reaction forces for needle level 2 and an OD of 50μm.

For the levels 2, 3 and 4, there is a good agreement between experiments and simulations concerning the vertical force as shown in Figure II.31 (b), (c) and (d). There is an increase of the vertical reaction force with increasing the over travel as expected because increasing the over travel increases the force applied to the pad and thus the reaction force measured at its surface. Hence, for those geometries, the assumptions established before are correct.

For the level 1 geometry of probe, the match is not as good as in the previous cases (see Figure II.31 (a)). This variation is probably due to the fact that this level has a slightly different geometry from the other probes. Hence, for this geometry, the assumptions have to be re-evaluated to better fit the experimental results. The assumption the most geometry dependant is the Young's modulus of the effective volume. The vertical reaction force simulated for the Figure II.31 (a) is 16mN for an OD of 75 μ m. The Young's modulus of the effective volume to fit with experiment has to be increased (see Figure II.24), and an elastic modulus of 300GPa matches the experimental value. This elastic modulus is valid for all the overdrives. With this new value of Young's modulus for the level 1 geometry of probe, the simulations and experiments are in good agreement.

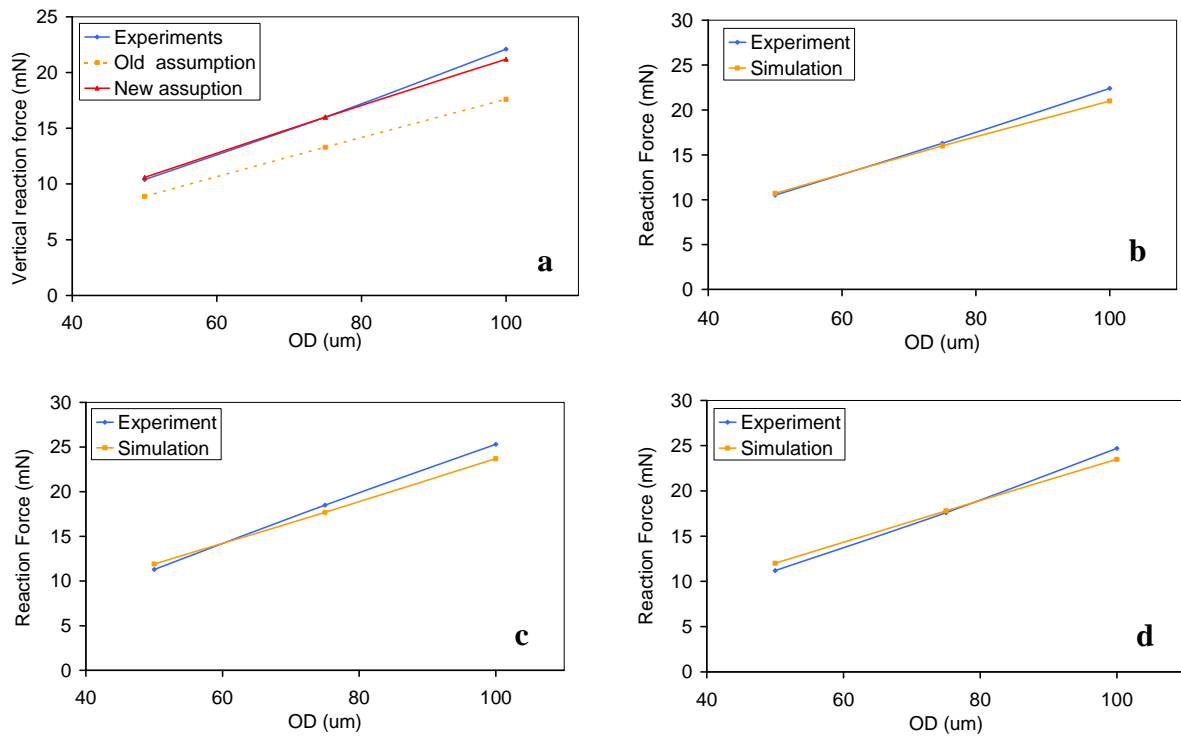


Figure II.31. Vertical reaction force for various overdrives for needle geometries: (a) level 1, (b) level 2, (c) level 3 and (d) level 4.

The Figure II.32 shows the comparison between the experimentally observed scrub marks left by the probe and the ones obtained by simulations. The measurement of the maximum length of the marks is performed with the help of an optical microscope for the experiments. For the modelling, the lateral displacement of the probe is added to its tip's diameter (i.e. 19 μ m). This assumption concerning the model leads to an over evaluation of the actual scrub mark, which underlines the complexity of the contact and the root causes of this over evaluation are later explained. However, the main point is that measured and simulated scrub marks are still in the same range of values, that is to say between 25 μ m and 49 μ m for the experiments and 31 μ m and 61 μ m for the modelled ones.

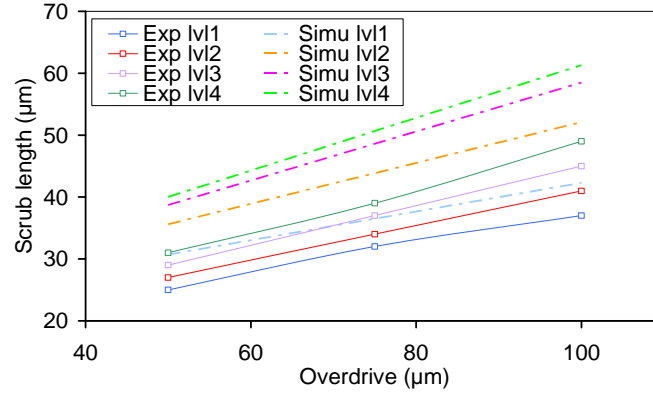


Figure II.32. Experimental and numerical scrub marks for various overdrives.

Finally, a good agreement is found for the modelling of the probing on CMOS 120nm with H9 probe card. Some effects such as the increase of the overdrive on the vertical and lateral reaction forces are noticed, and further investigation are required to see if it can also be related to higher damages in the pad.

III.B.1.c) Energetic considerations

The simulated energies have to be checked because energies such as hourglass (see section III.A.4.e) for additional details) are purely numerical and can invalidate the results. The total energy of the system is the following:

$$E_{total} = E_{conservative} + E_{dissipated} \quad (II. 17)$$

where the conservative energies are the internal and the kinetic energies and the dissipated energies are the sliding and the hourglass energies. More precisely in our case:

$$E_{total} = E_{internal} + E_{kinetic} + E_{sliding} + E_{hourglass} \quad (II. 18)$$

where $E_{sliding}$, which is linked to the energy loss due to friction between the pad and the probe, is equal to:

$$E_{sliding} = \mu \int_{space} F_{vert}(x) * dx \quad (II. 19)$$

where μ is the friction factor and $F_{vert}(x)$, the vertical force that depends on the lateral position of the probe on the pad.

For the EWS model, since the probe is light and has a small velocity, the kinetic energy is negligible in front of the other energies. As shown in Figure II.33, the total energy behaves like the overdrive applied with loading, holding and unloading stages. The internal energy is the major contributor to the total energy. The sliding energy increases as described by the equation (II. 19). The hourglass energy is very low (less than 5% of the internal energy) and has the same shape than the internal energy. Hence, the use of reduced elements is not detrimental to the simulation here. The total energy at the end of the probing is not null due to the energy loss linked to the friction during the probing.

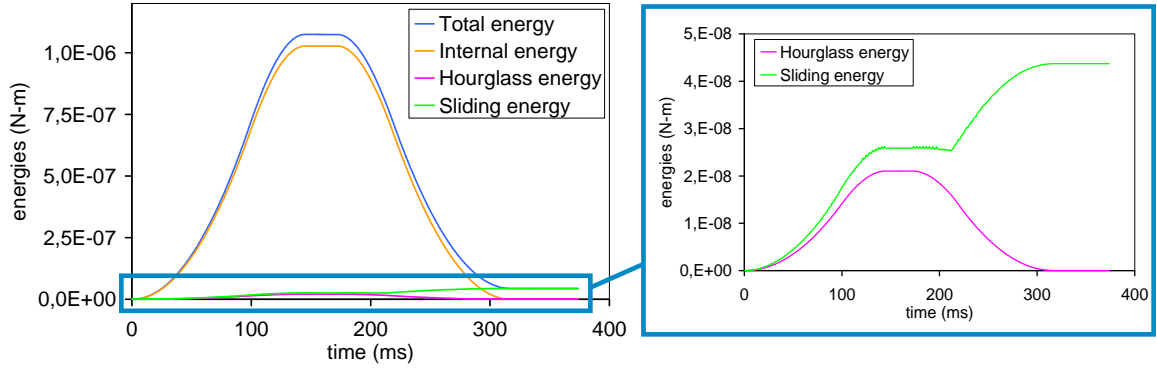


Figure II.33. Evolution of the total, internal, sliding and hourglass energy during a typical EWS simulation.

III.B.2. Experiment and calibration for CMOS 45nm pad model

The probing on the CMOS 45nm technology node is made with Helium+ cantilever probe card, whose needle geometry is shown in Figure II.34. Hence, similarly to the calibration with the H9 probe card, dedicated experiments are required to get a reliable model.

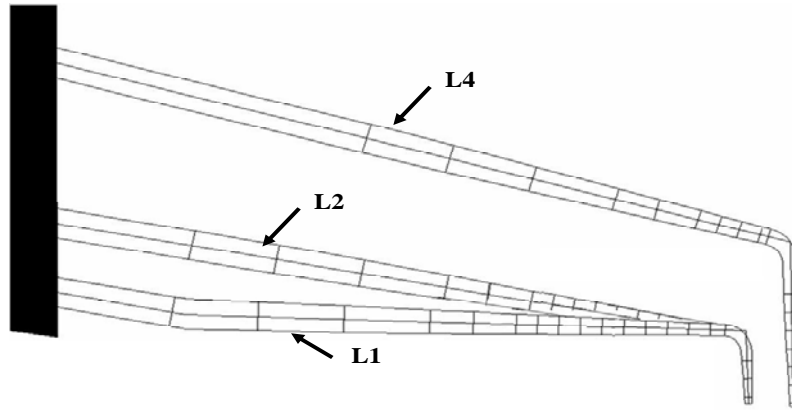


Figure II.34. Three of the geometries of probe and probe tip studied for the Helium+ probe card.

III.B.2.a) Experimental measurements of the reaction forces

For the Helium+ probe card, three of the four needles are tested (level 1, 2 and 4). The overdrive is fixed at 90μm and the number of touchdowns varies from 1 to 8. The scrub mark resulting from one of the test is shown in Figure II.35. The lengths and widths of the damaged areas are in the range of 21μm and 16μm respectively. The length of the marks is much smaller than these with the previous probe cards while their width is close.

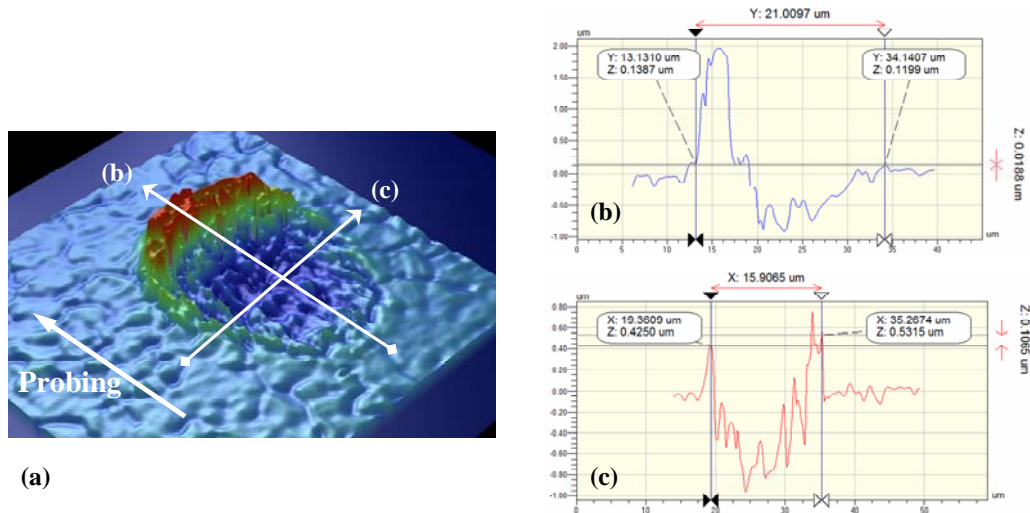


Figure II.35. Optical profilometry view of a probe mark (a) and profiles of the mark (b and c).

The reaction forces measured are depicted in the Figure II.36. The ranking according to the force, whether vertical or lateral, applied to the pad by each needle is deduced:

$$Level1 < Level2 < Level4 \quad (II. 20)$$

It has to be noted that the ranking obtained for that probe card is the same that the one of the H9 probe card (II. 16). Moreover, the lateral force induced by the level 1 geometry is lower than the one from the two other needles. On the contrary, the vertical force induced by the level 4 geometry is much higher than these of the two other needles. This induces higher differences for the friction coefficient that depends also on the needle geometry in that case, and varies from 0.48 (level 2) to 0.68 (level 4). The friction coefficient found for this probe card is much higher than the one obtained with the previously considered probe card (0.096).

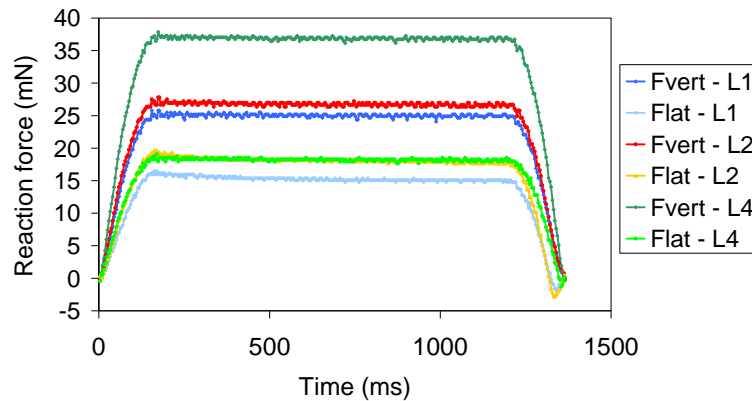


Figure II.36. Vertical and lateral reaction force for the various needles for the eighth touchdown.

The effect of the number of touchdown on the reaction force is depicted in Figure II.37. Both vertical and lateral reaction forces increase when the number of touchdown increases. The ratio of forces between the first and the eighth touchdown reaches 0.92 and 0.71 respectively for the vertical and the lateral reaction force. Hence, only a small increase of the vertical force is observed. For the lateral force, this effect is much more important. However, in both cases, after about five touchdowns, the maximum value is reached.

The reasons for the increase of the vertical and lateral reaction forces may have two sources. The first one is the formation of the pile-up that is located in front of the mark (refer to the green and red areas on the Figure II.35) and the second one is the work hardening of the

aluminium layer during the successive tests. These root causes are later more widely explained.

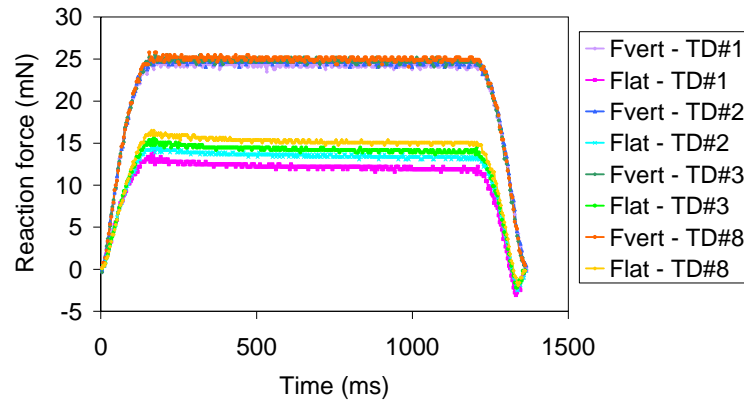


Figure II.37. Vertical and lateral reaction forces for level 1 needle geometry.

III.B.2.b) Model calibration and issues

The simulation related to the experimental conditions is performed (same OD and number of touchdowns). However, as previously mentioned, the friction coefficient is much higher than during the calibration of the H9 probe card. The length of the scrub is smaller as well. Hence, the friction coefficient has to be fixed at 0.58 to try to mimic the experiments.

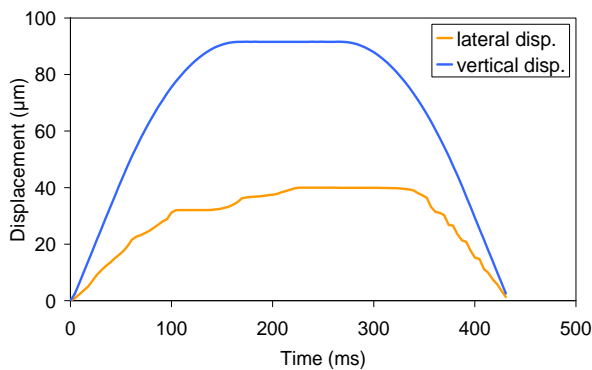


Figure II.38. Vertical and lateral displacement of the probe tip during one touchdown.

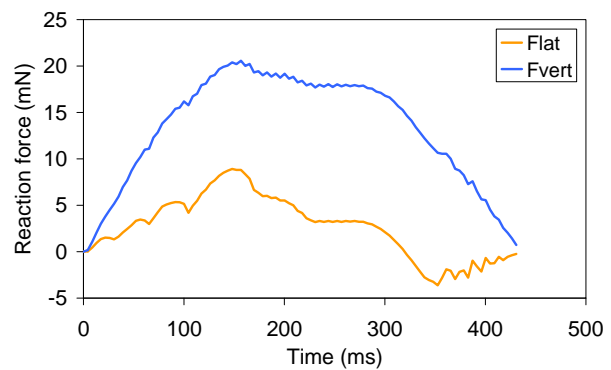


Figure II.39. Vertical and lateral reaction forces during one touchdown.

The Figure II.38 shows the displacement of the probe tip during one touchdown. The vertical displacement evolves as expected and shows no contact loss. The behaviour of the lateral displacement shows a plateau during the loading stage due to the increase of the friction coefficient. A sort of slip-stick motion thus appears. Moreover, due to the higher friction coefficient, the delay between the beginning of the unloading and the beginning of the lateral motion of the tip is very important. Both the vertical and lateral reaction forces no longer evolve linearly with the overdrive as shown in Figure II.39. This is also linked to the high friction coefficient. The quick variations of the vertical and lateral reaction forces have different sources. The first one is the slip-stick motion, as mentioned here above. A loss and regain of the contact between the needle and the pad and vibrations in the body of the needle can be related to jumps of the forces as well. Finally, the concept of friction coefficient between the pad and the tip is not valid anymore in that particular case.

In the numerical model, the friction coefficient is computed by the relation [Ansys LSDyna, rel. 11]:

$$\mu = \mu_d + (\mu_s - \mu_d)e^{-cd(v)} \quad (\text{II. 21})$$

where μ_s is the static friction coefficient, μ_d is the dynamic friction coefficient, cd is the decay coefficient. Additionally, the friction coefficient depends on the relative velocity v between the slave and the master segments.

In the literature, many papers [Reid, 2004; Hashiguchi, 2005; Hashiguchi, 2008; Peilleux, 2008] describe issues encountered while using explicit models with high friction coefficients, or describe some new algorithms to try to solve related issues. However, these new algorithms are not implemented yet in the simulation software. In particular, Reid et al. underlines among other meshing and solving conditions (use of fully integrated elements, fine meshing and low time steps). The use of the penalty contact algorithm, which is different from the actual physical friction phenomenon, is emphasized as well. All these effects lead to very long computing time additionally to the point aforementioned. Hence, a dedicated modelling study would be required to calibrate the model and in particular reproduce properly the lateral displacement and induced forces of the tip.

III.B.3. Modal analysis of the probe behaviour

The modal analysis is performed for the cantilever H9 probe card. The main purpose while performing a modal analysis is to evaluate the response of the geometrical design to dynamic loading. Such an analysis gives the natural frequencies and the shapes of the modes. It can also be used to help calculating the solution control (i.e. mass scaling) for the dynamic analysis.

III.B.3.a) Problem description and model extraction

Assuming that the model is free of damping and vibrations, the equation (II. 11) becomes:

$$[M]\{\ddot{u}\} + [K]\{u\} = \{0\} \quad (\text{II. 22})$$

Moreover, assuming harmonic motion (i.e. $u = U \sin(\omega t)$), (II. 22) can be rewritten:

$$([K] - \omega^2[M])\{u\} = \{0\} \quad (\text{II. 23})$$

where the matrixes $[K]$ and $[M]$ are large, sparse and symmetric positive.

The roots of the equations are ω_i^2 , the eigenvalues, where i ranges from 1 to the number of degrees of freedom, and corresponding eigenvectors are $\{u\}_i$. The square roots of the eigenvalues ω_i are the structure's natural circular frequencies. Hence, the natural frequencies f_i are obtained using $f_i = \omega_i / 2\pi$. The eigenvectors $\{u\}_i$ represent the mode shapes, that are the shapes of the structure while vibrating at frequency f_i .

In order to extract these eigenvectors and eigenvalues, the block Lanczos algorithm is chosen (see also [Ansys LSDyna, rel. 11]). The theoretical basis of the solver is described by Grimes et al. [Grimes, 1994]. This method uses blocks of vectors and the Lanczos iteration to obtain the higher mode shapes and related frequencies (the 4 principal modes in this case).

III.B.3.b) Modal analysis

For the model, which contains only the needle, a vertical motion is applied to the end of the tip that can move either freely or not. Thus, five cases are investigated, as depicted in Table II.4, in order to evaluate the effects of the friction and the mass scaling.

Case	1	2	3	4	5
Lateral displacement of the probe tip	Free	Free	Free	Free	Locked
Mass scaling	None	No mass	Mass of the tip *10	Mass of the tip *10 ³	None

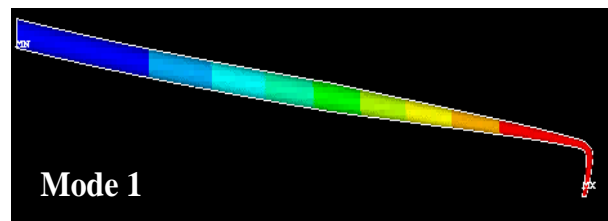
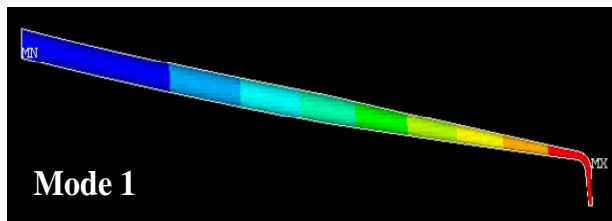
Table II.4. Simulated cases investigated to check the effects of the mass scaling and friction

While the lateral displacement of the probe tip is set as free, it corresponds to a contact with pure sliding that is to say with a null friction coefficient (or scrubbing). On the contrary, when the displacement of the tip is locked, the model assumed underlies no sliding or sticking (or stubbing). The mass scaling is only applied to the tip since it contains the smallest elements (that control the time step as described in section III.A.4.b)) to get a good management of the contact. The Table II.5 shows the results of the modal analysis, for the four principal vibration modes. The effect of the mass scaling is shown to have little effect on frequency of the first vibration mode. Yet, this decrease of the frequency of oscillations becomes very important for higher vibration modes and the frequency of the fourth mode is 60% smaller than the probe with actual mass (cases 1 to 4). For our model, the mass of the whole probe is 0.8949mg and the mass of the tip is $0.96 \cdot 10^{-5}$ mg (only half of the model is considered for simulations). The effect of the mass scaling has to be taken into consideration when higher vibration modes are considered as part of motion of the probe, i.e. while dynamics of the needle can not be neglected (see next chapter for the effect of the velocity during probing).

		Case				
		1	2	3	4	5
Frequencies (Hz)	Mode 1	13180	13181	13175	12594	14746
	Mode 2	58523	58536	58409	46502	63709
	Mode 3	117810	117331	117331	77162	129613
	Mode 4	205244	205443	203406	122569	228713

Table II.5. Oscillating frequencies for the 4 principal modes and the 5 considered cases.

The effect of the friction (slipping and sticking options) is also considered in the study. The stubbing is related to higher frequencies of the vibration modes. Moreover, the shapes of the vibration modes are also changed by locking the end of the tip. The three major vibration modes for the scrubbing and the stubbing cases are shown in Figure II.40 and Figure II.41. Hence, the sliding, controlled by the friction coefficient is shown once more to have a huge effect on the deformation and vibration frequencies of the needle.



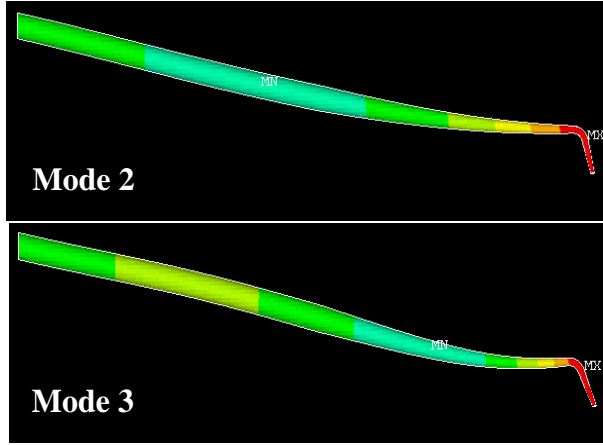


Figure II.40. Principal mode shapes for case 1 (free probe tip).

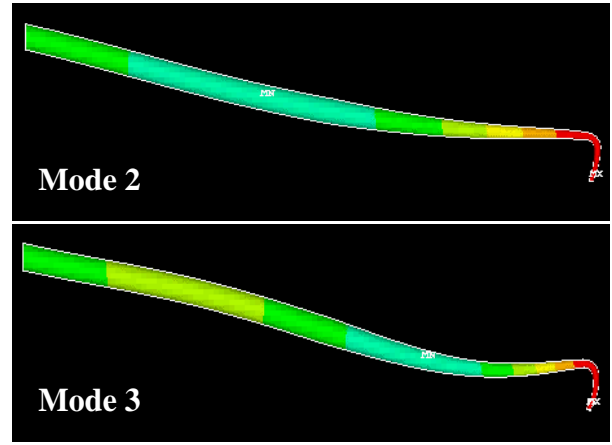


Figure II.41. Principal mode shapes for case 5 (locked probe tip).

In the literature, a modal analysis has been performed by Schmadlak et al. [Schmadlak, 2007] on a cantilever needle. For the scrubbing and the stubbing first modes, they obtain the same needle deformation shapes. However, they get different frequencies, about 4 times bigger than these of the current work. The reason is the distinction in the geometrical model and more precisely the size and the mass of the needle.

III.B.3.c) Comparison with explicit model

In order to compare the vibration modes obtained during the modal analysis and these of the explicit model, a specific loading is performed, as shown in Figure II.42. Firstly, the pad is raised to reach an overdrive of 100μm, then the pad is hold at a plateau in order to avoid vibration due to the loading, and finally, the pad is removed, and the free, non-damped oscillation can start. A mass scaling of $5 \cdot 10^{-7}$ is selected, which correspond to an increase of the mass by a factor of $4.6 \cdot 10^3$. The analysis of these oscillations is made using the Fast Fourier Transform (FFT).

The Fourier analysis is the decomposition of a function in terms of sinusoidal functions of different frequencies that can be recombined to obtain the original function. The recombination process is the Fourier synthesis. The result of the decomposition is the amplitude and the phase, which are imparted to each basis frequency in the reconstruction. The main interest of the FFT is its reduced computation time, due to a lower number of operations. A fast Fourier transform (FFT) is an algorithm to compute the discrete Fourier transform (DFT) and its inverse. The DFT is defined by the formula:

$$X_k = \sum_{n=0}^{N-1} x_n e^{-\frac{2\pi i}{N} nk} \quad (\text{II. 24})$$

where x_0, \dots, x_{N-1} are complex number, and k varies from 0 to $N-1$.

As depicted in the Figure II.43, the frequencies corresponding to the first and the second modes are respectively 12.5 kHz and 44 kHz, which are in good agreement with the values obtained during the modal analysis with mass of the tip multiplied by 1000 (see Table II.5). A third frequency is found at 56 kHz. This frequency cannot be related a vibration mode¹. The amplitude of the first vibration mode is much higher than other ones (about 20 times), as a

¹The small difference of mass scaling between the modal analysis (mass*1000) and the dynamic model (mass*4600) is expected to reduce the frequency of the third vibration, but not in a such important range (77 kHz versus 56 kHz) .

consequence, the first vibration mode is the main contributor to the deformation of the needle during the simulation.

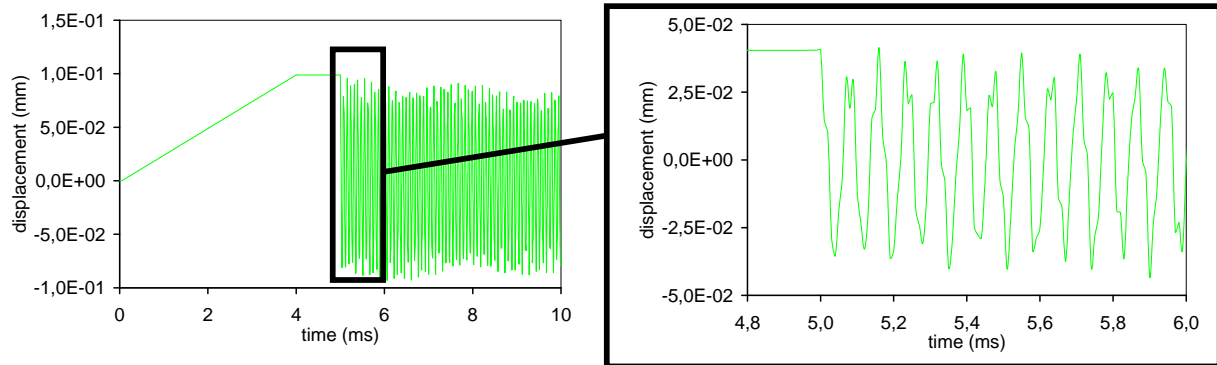


Figure II.42. Loading conditions for the modal analysis and zoom at the initial oscillations.

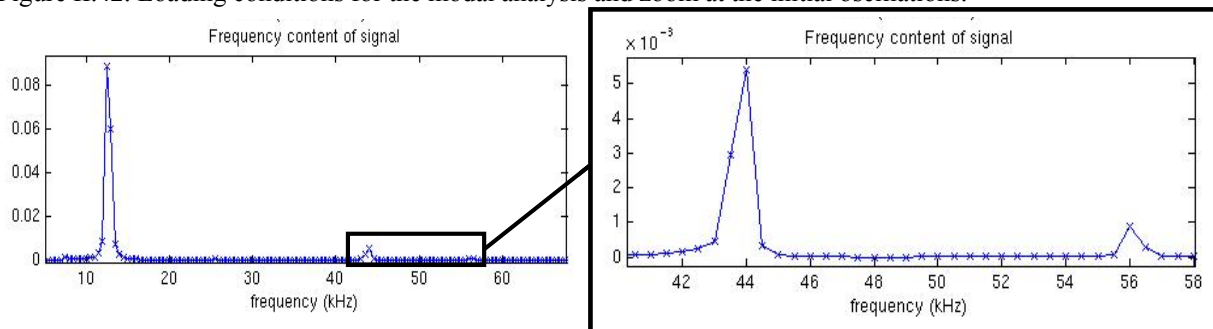


Figure II.43. Frequencies of the 3 major vibration modes from explicit modelling.

Finally, the first vibration mode of the needle has the major role in the deformation of the probe. As a consequence, the mass scaling has limited effect on the probe motion when the needle dynamics is not too important.

IV. Conclusions

In this chapter, the various analysis techniques to observe the pad such as the FIB milling followed by SEM observations or the optical profilometry have been described. The advantages and drawbacks of each of these techniques have also been underlined. These imaging tools enable us to improve the understanding of the damages and failures undergone by the pad during the probing and wire bonding operations, providing data like the profile of the scrub mark or the localization of the cracks under the surface.

In a second part, the principle of the nanoindentation is described. This experimental technique is set up to evaluate the mechanical robustness of the pad structure and to reproduce the loading conditions and failures related to EWS and assembly steps. Then, the Oliver and Pharr model and its limitations are introduced. The Berkovich and cube corner tips that are later used during testing are presented as well.

In a third section, the numerical model to mimic EWS is described, starting with the description of the geometries of the needles evaluated. Then, the material properties used for the needle and the pad are justified. The physical (e.g. damping, effective volume...) and numerical (e.g. mass scaling, reduced integration...) parameters leading to inaccuracy and even errors in the modelling results are checked not to influence them in a large range. In order to build a model valid for several technology nodes (CMOS 45nm to CMOS 120nm) and to fairly reproduce the probing, the calibration of two cantilever probe cards is described. For the first one (H9 for CMOS 120nm testing), a good agreement is obtained between the

experiments dedicated to the calibration and the model. As a consequence, this model is later widely used. For the second one (Helium+ for CMOS 45nm probing), the numerical instability due to the high friction coefficient does not able to obtain reliable results. Finally, a modal analysis (for H9 probe card) has been performed to investigate the effects of mass scaling and sliding. The former has been proved to have little effect while the dynamical effects are not involved, and the later is shown to modify in a large range the deformation of the needle.

In the next chapters, the imaging and the experimental and numerical tools presented and set up in this chapter are used to evaluate various tests and design conditions. Additional refinements and indicators to improve the nanoindentation and the finite element model capabilities to compare the mechanical robustness of pads will be defined. These techniques will be faced to actual tests and correlations, but also few limitations of the tools will be described.

References

[Ansys LSDyna, rel. 11] Ansys LSDyna user manual, release 11.

[Bolshakov, 1998] Bolshakov, A. and Pharr, G.M., Influences of pileup on the measurement of mechanical properties by load and depth sensing indentation techniques, *Journal of materials research*, vol. 13, no. 4, April 1998, pp. 1049-1058.

[Boussinesq, 1885] Boussinesq, J., *Application des potentiels à l'étude de l'équilibre et du mouvement des solides élastiques*, Ed. Gauthier-Villard, Paris, 1885

[Bulychev, 1975] Bulychev, S.I., Alekhin, V.P., Shorshorov, M.K., Ternovskii, A.P. and Shnyrev, G.D., Determining Young's modulus from the indenter penetration diagram, *Zavod. lab.*, vol. 4, no. 9, 1975, pp. 1137-1140.

[Cacho, 2007] Cacho, F., Vettori, R., Galbiati, G., Thiery, G., Fiori, V., Ainaoui, M., Delpech, P., Tavernier, C. and Jaouen, H., Study of probing test for electrical wafer sort application: dynamical simulation and experimental measurement, *IEEE 9th Electronics packaging technology conference*, 2007, pp.391-395.

[Chutoba, 2001] Chudoba, T. and Richter, F., Investigation of creep behaviour under load during indentation experiments and its influence on hardness and modulus results, *Surface and coatings technology* 148, 2001, pp. 191–198.

[Doerner, 1986] Doerner, M.F. and Nix, W.D., A method for interpreting the data from depth-sensing indentation instruments, *Journal of material research*, Vol. 1, Issue 4, July/August 1986, pp. 601-609.

[Fiori, 2005] Fiori, V. and Orain, S., A multi scale finite element methodology to evaluate wire bond pad architectures, *IEEE International conference on thermal, mechanical and multi-physics simulation experiments in microelectronics and micro-systems*, Eurosime, 2005, pp. 648-655.

[Gallois-Garreignot, 2009] Gallois-Garreignot, S., Chave, F., Gonchond, J.P., Gautheron, B., Fiori, V. and Nelias, D., Thin films interfacial adhesion characterization by cross-sectional nanoindentation: application to pad structures, *IEEE 11th Electronics packaging technology conference*, 2009, pp. 947-952.

[Grimes, 1994] Grimes, R., Lewis, J.G. and Simon, H., A shifted block Lanczos algorithm for solving sparse symmetric generalized eigenproblems, *SIAM Journal matrix analysis application*, vol. 15, no. 1, 1994, pp. 228-272.

[Hashiguchi, 2005] Hashiguchi, K., Ozaki, S. and Okayasu, T., Unconventional friction theory based on the subloading surface concept, *International journal of solids and structures* 42, 2005, pp. 1705–1727.

[Hashiguchi, 2008] Hashiguchi, K. and Ozaki, S., Constitutive equation for friction with transition from static to kinetic friction and recovery of static friction, *International journal of plasticity* 24, 2008, pp. 2102–2124.

[Hay, 1999] Hay, J.C., Bolshakov, A. and Pharr, G.M. Critical examination of the fundamental relations used in the analysis of nanoindentation data, *Journal of material research*, vol. 14, no. 6, 1999, pp. 2296-2305.

[Hay, 2000] Hay, J.L. and Pharr, G.M., Instrumented indentation testing, *Material park, ASM international*, 2000, pp. 232-243.

[Hertz, 1881] H. Hertz, Über die Berührung fester elastischer Körper, Journal für die reine und angewandte Mathematik 92, 1881, pp. 156-171.

[Humphreys, 2001] Humphreys, F.J., Grain and subgrain characterization by electron backscattered diffraction, Journal of materials science, vol. 36, 2001, pp. 3833-3854.

[King, 1987] King, R.B., Elastic analysis of some punch problems for a layered medium, International journal of solids structures, vol. 23, no. 12, 1987, pp. 1657-1664.

[Li, 2002] Li, X. and Bhushan, B., A review of nanoindentation continuous stiffness measurement technique and its applications, Materials characterization, vol. 48, issue 1, February 2002, pp. 11-36.

[Liu, 2006] Liu, D.S. and Shih, M.K., Experimental method and FE simulation model for evaluation of wafer probing parameters, Microelectronics journal 37, 2006, pp. 871-883.

[Oliver, 1992] Oliver, W.C. and Pharr, G.M., An improved technique for determining hardness and elastic modulus using load and displacement sensing indentation experiments, Journal of materials research 1992, vol. 7, issue 6, pp. 1564-1583.

[Oliver, 2004] Oliver, W.C. and Pharr, G.M., Measurement of hardness and elastic modulus by instrumented indentation: advances in understanding and refinements to methodology, Journal of materials research, vol. 19, no. 1, January 2004, pp. 3-20.

[Peillex, 2008] Peillex, G., Baillet, L. and Berthier, Y., Homogenization in non-linear dynamics due to frictional contact, International journal of solids and structures 45, 2008, pp. 2451-2469.

[Reid, 2004] Reid, J.D. and Hiser, N.R., Friction modelling between solid elements, International journal of crashworthiness, vol. 9, no. 1, January 2004, pp. 65-72.

[Sanchez, 1999] Sanchez J. M., El-Mansy, S., Sun, B., Scherban, T., Fang, N., Pantus, D., Ford, W., Elizalde, M.R., Martinez-Esnaola, J.M., Martinez-Meizoso, A., Gil-Sevillano, J., Fuentes, M. and Maiz, J., Cross-Sectional Nanoindentation: a new technique for thin film interfacial adhesion characterization, Acta Materialia vol. 47, no. 17, 1999, pp. 4405-4413.

[Schmadlak, 2007] Schmadlak, I. and Hauck, T., Simulation of wafer probing process considering probe needle dynamics, IEEE International conference on thermal, mechanical and multi-physics simulation experiments in microelectronics and micro-systems, Eurosime, 2007, pp. 1-5.

[Sneddon, 1965] Sneddon, I.N., The relation between load and penetration in the axisymmetric Boussinesq problem for a punch of arbitrary profile, Int. J. Eng. Sci. 3, 1965, pp. 47-57.

[Tabor, 1948] Tabor, D., A simple theory of static and dynamic hardness, Proceeding from research society, London, ser. A 192, 1948, pp. 247.

[Wu, 2006] Wu, F. and Gonchond, J.-P., Optimisation de la technique de nanoindentation pour la caractérisation des propriétés mécaniques des oxydes basse permittivité, Université Joseph Fourier, Grenoble, France, Rapport de stage de fin d'étude, 2006.

Chapter 3:

Test and assembly critical parameters on the appearance of mechanical issues in pads

Outline

I.	EWS tests parameters effect on the crack occurrence.....	82 -
I.A.	Dynamical effects during probing.....	82 -
I.A.1.	Chuck velocity.....	82 -
I.A.2.	Chuck vibrations	84 -
I.B.	Test parameters effect: overdrive and number of passes	84 -
I.B.1.	Observation and measurements of the probe mark	85 -
I.B.2.	Pad integrity results and observations	88 -
I.B.3.	Finite element model results and limitations.....	89 -
I.B.4.	Discussion	90 -
II.	Gold bonding: Influence of the process parameters.....	91 -
II.A.	Gold bonding with 25 μ m diameter wires.....	91 -
II.A.1.	First bonding Design of Experiments.....	92 -
II.A.2.	Second wire bonding DOE.....	95 -
II.A.3.	Discussion	96 -
II.B.	Gold bonding with 20 μ m diameter wires.....	97 -
II.B.1.	Bonding parameters effects on mashed ball dimension	97 -
II.B.2.	Wire pull tests results	98 -
II.B.3.	Ball adhesion on pads.....	98 -
II.B.4.	Discussion	99 -
II.C.	Effect of the thermal aging.....	99 -
II.C.1.	Wire pull tests results	100 -
II.C.2.	Ball adhesion on pads.....	100 -
II.D.	Discussion	101 -
III.	Copper bonding: effect of the process parameters	101 -
III.A.	Initial bonding observations and copper ball measurements	102 -
III.A.1.	Aluminium pad splash.....	102 -
III.A.2.	Mashed ball measurements	103 -
III.B.	Wire pull tests results	103 -
III.C.	Shear tests results	105 -
III.C.1.	Effect of the shear test parameters on forces.....	105 -
III.C.2.	Effect of the contact force and the bonding power	106 -
III.D.	Discussion	108 -
III.D.1.	Comparison between gold and copper bonding	108 -
III.D.2.	Weight of the process and design parameters	109 -
IV.	Conclusions	110 -
References	111 -

In the previous chapters, the electrical wafer sort and the wire bonding have been described and shown to be able to generate failures in the pads. Additionally, the common pad integrity analyses to observe failures induced by probing and the wire pull and shear tests have been presented. Then, some experimental and numerical tools and techniques have been set up to reproduce the actual tests.

In this chapter, the test and assembly parameters are investigated through experimental and numerical designs of experiments in order to find out which ones are detrimental to the pad structure. In that frame, the target is to induce failures (cracks and peeling) in the pad to compare them, even if in the manufacturing of an I.C., these must be avoided. For the EWS, the dynamical effects such as the velocity of the chuck or its vibrations are considered with the help of the numerical model. Then, the scrub marks left by the probing test on aluminium are measured. After that, the overdrive and the number of passes, which are EWS test parameters, are investigated through both actual tests and numerical studies. In a second part, the bonding parameters are weighted for both gold and copper bonding. To that purpose, ball size measurements, wire pull and shear tests are performed, which allow the evaluation of the bonding resistance. Finally, all the tests and process parameters are compared with the design ones.

I. EWS tests parameters effect on the crack occurrence

I.A. Dynamical effects during probing

I.A.1. Chuck velocity

I.A.1.a) Problem position

The velocity of the chuck during the impact between the needle and the pad is suspected by the ST Grenoble EWS team to be responsible for the appearance of the failures. The effect of this parameter can be evaluated during actual tests, using various velocities and performing pad integrity tests. Some MEMS structures can be used as well to measure the forces undergone by the pad. Both techniques are quite long and expensive to set up and a simple way to check this effect is through modelling. The velocity of the chuck during actual EWS is between 1mm/s and 20 mm/s. Hence, in the following study, chuck velocities from 4mm/s to 4000mm/s are evaluated.

I.A.1.b) Numerical results

Only the loading stage is considered here, and the overdrive is fixed at 100 μ m. For comparison purpose, the time is normalized between the beginning and the end of the loading (time=1 corresponds to the time when the chuck reaches the 100 μ m). The study is based on the vertical reaction force since it allows a good evaluation of the contact between the pad and the needle.

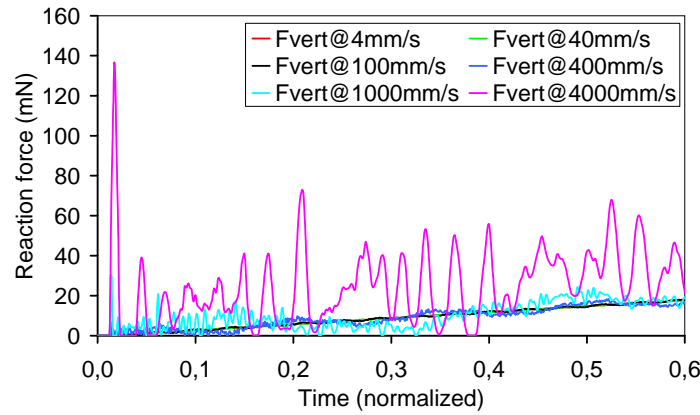


Figure III.1. Vertical reaction force versus normalized time for the various velocities investigated. The end of the loading ($t=1$) is reached at the maximum overdrive.

The vertical reaction forces for the various chuck velocities considered are shown in the Figure III.1. At the end of the loading, the maximum vertical reaction force obtained is 29.3mN. For $v=4000\text{mm/s}$, at the impact time, the reaction force reaches 136mN. Hence, it is much higher than the force at the end of the loading, and issues related to the impact on the pad may occur. The threshold for the appearance of this dynamical effect is $v=1000\text{mm/s}$. For such a velocity, both dynamic and static effects must be considered and the forces undergone by the pad are in the same range ($F_{\text{vert}}=30.1\text{mN}$ for the impact).

For velocities bellow 400mm/s, the force during the impact becomes negligible in front of the force at the end of the loading and even not detectable bellow 100mm/s, as depicted by the Figure III.2.

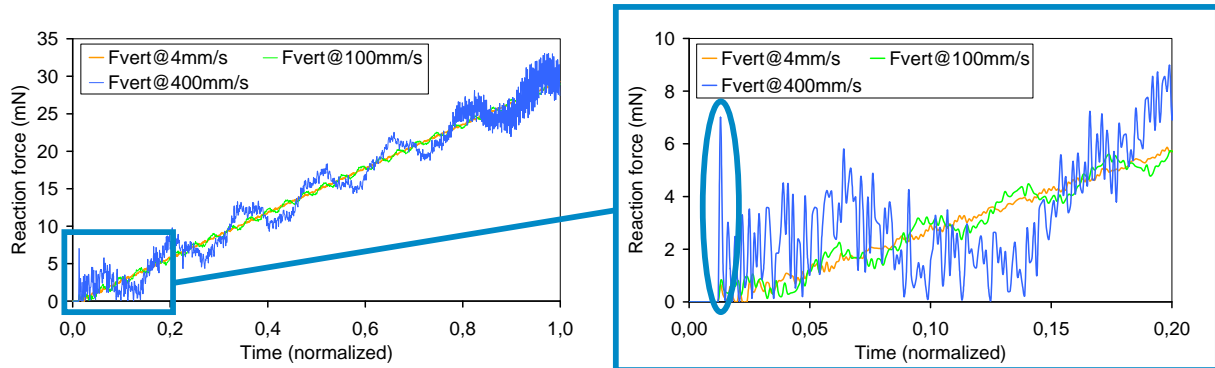


Figure III.2. Vertical reaction force versus normalized time for the lowest velocities.

In the literature, the effect of the velocity has been investigated by Cacho et al. [Cacho, 2007] and by Schmadlak et al. [Schmadlak, 2007], and similar results have been obtained. The former have shown that for low velocities (below 16mm/s), the impact does not induce any oscillation in the vertical reaction force. The latter have demonstrated that very small oscillations were obtained for $v=35\text{mm/s}$ (the impact force reaches 14mN while the maximum force reaches 140mN, i.e. about 10%), and as aforementioned, the geometrical design of the probe is slightly different that may explain the variation observed here. Hence, in both papers, it is deduced that the dynamical effects do not need to be considered.

Finally, from the results of the numerical model, the current velocity during the probing process may not induce damages in the pad and a threshold value for their appearance is given. This result is later confirmed with the crack location (see next chapter). Increasing the chuck velocity is related to a reduced time to test the wafers and consequently to a technology speed up.

I.A.2. Chuck vibrations

The vibrations of the chuck at the end of the loading stage can induce additional non expected stress in the structure. These are caused by the inertia of the chuck that cannot stop immediately at the end of the loading stage and whose position oscillates around the fixed overdrive. This issue has been suspected by EWS team at ST Grenoble due to differences in failure analyses after probing on the same samples but different probers. A dedicated experiment is required to evaluate this effect for overdrives of 50 μm and 90 μm . Thus a laser is positioned on the top of the prober, in order to measure the position of the chuck during the EWS test. The schematic of the apparatus is presented in Figure III.3.

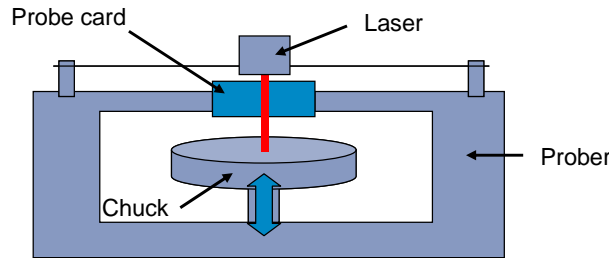


Figure III.3. Schematic apparatus set up for the measurement of the vertical displacement of the chuck during probing.

The measured position of the chuck during the first touchdown is shown in Figure III.4. Some small oscillations are found at the end of the loading and unloading stages, reaching less than 1.3 μm , whatever the overdrive.

Theses displacement profiles are implemented in the finite element model and the reaction forces are analyzed. Thus, the increase of the reaction forces is lower than 0.2mN whatever the OD, which is less than 1% of the maximum vertical force undergone by the pad. As a consequence, the vibrations of the chuck during the probing can be neglected, and cannot explain the differences observed after the pad integrity analyses.

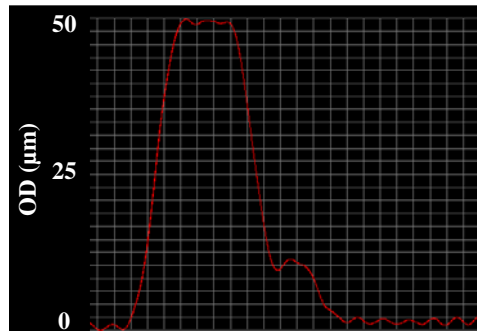


Figure III.4. Laser measurement of the position of the chuck during one touchdown (OD=50 μm).

Finally, considering the tests parameters currently used, the velocity of the chuck and its inertial effects can be neglected during further analysis, but they have to be kept in mind since they might be the causes of damages for the next technology nodes with more fragile structures containing low-k oxides.

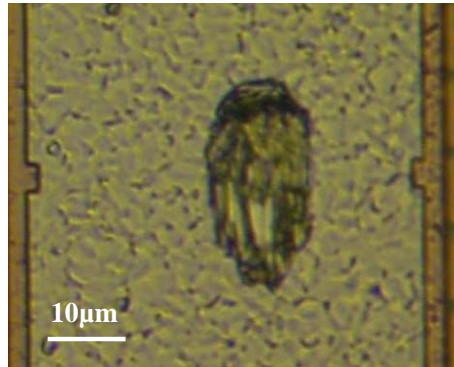
I.B. Test parameters effect: overdrive and number of passes

The test parameters of EWS investigated in this study are the number of passes and the overdrive. Either 5 or 10 passes with overdrives of 50 μm and 90 μm have been performed on patterned wafer CMOS 120nm technology node.

I.B.1. Observation and measurements of the probe mark

I.B.1.a) Scrub mark measurements

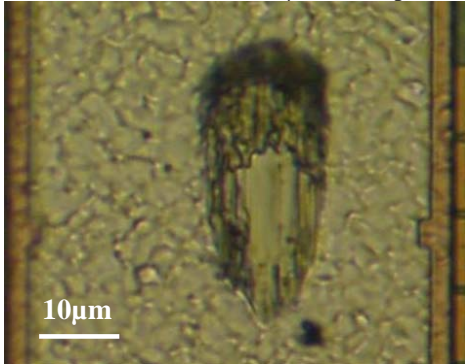
After being probed, the scrub marks are observed by means of an optical microscope. The Figure III.5 shows top views of the marks on aluminium. Scrub marks have an ovoid shape that is related to the slip of a cylinder on a flat surface.



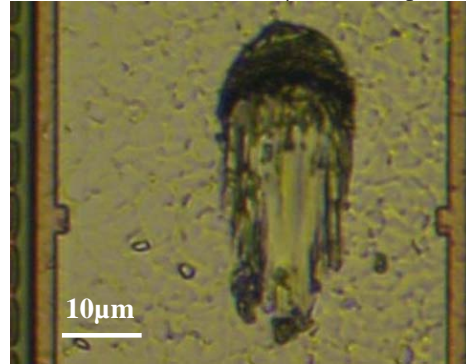
(a) Scrub mark for OD 50μm and 5 passes



(b) Scrub mark for OD 50μm and 10 passes



(c) Scrub mark for OD 90μm and 5 passes



(d) Scrub mark for OD 90μm and 10 passes

Figure III.5. Optical microscopy top views of scrub marks for the various probing conditions.

Thus, the marks are measured and the average over 20 marks is made. The averages and standard deviations are gathered in the Table III.1. It is found that increasing the overdrive from 50μm to 90μm induces an increase of both the length and the width of the scrub mark. The former is due to a bigger bending of the probe's body since the chuck is raised higher and the tip's end is fixed. The radius of the probe tip is 15μm, and the radius of curvature is 2μm, thus the increase of the width may be due to a deeper penetration of the probe into the aluminium layer.

	Average length (μm)	Average width (μm)
OD 50μm – 5 passes	27.29 +/- 2.15	13.89 +/- 1.02
OD 50μm – 10 passes	27.18 +/- 1.75	15.42 +/- 0.88
OD 90μm – 5 passes	40.39 +/- 2.89	17.67 +/- 0.84
OD 90μm – 10 passes	39.8 +/- 2.39	18.97 +/- 0.98

Table III.1. Scrub mark lengths and widths.

While increasing the number of passes from 5 to 10 for a given OD, the length of the probe mark does not increase. This result is correlated to the one observed during the calibration step while after 5 passes, the vertical and lateral reaction forces had reached their maximum. However, the widths of the marks increase by 1.53μm (+11%) and 1.3μm (+13%) respectively for the lowest and the highest OD. This increase is due to either an increase of

the contact area between the needle and the pad or to a small misalignment of the probe between the successive passes.

Moreover, the increase of the test parameters (mainly the overdrive) seems also to increase the damages and the removal of the aluminium from the top of the pad, as shown by the Figure III.5 (c) and (d), corroborating the assumption of a larger contact between needle and pad. The white areas in the centre of the mark are caused by a complete removal of the aluminium and thus a direct observation of the TaN/Ta layer.

In the literature, Chang et al. [Chang, 2010] have also observed an increase of both the scrub length and the scrub area while the overdrive increased, in the range between 35 μ m and 65 μ m, by means of actual tests. Liu et al. [Liu, 2006] have obtained from experiments and a finite element model that increasing the overdrive induces an increase of the scrub length, in the range 30 μ m to 100 μ m.

I.B.1.b) Aluminium damages after EWS

During the probing test, the aluminium top layer of the pad is plastically deformed. A 3D observation by means of optical profilometry of the surface is shown in the Figure III.6. The probe tip removes some aluminium from its original location, and pushes it away to form the pile up in front of the mark (red and green areas). As a result, in the centre of the mark, the aluminium is partially or completely removed and a depression is seen after the test (deep blue zone).

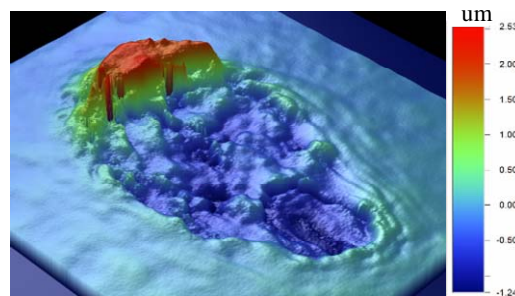


Figure III.6. 3D observation of the pad surface after EWS by optical profilometry.

The FIB/SEM technique allows observing the aluminium grains and reconstructing them in 3 dimensions. The Figure III.7 shows a SEM view of the cross section of the pad after FIB milling. The pile up and removed aluminium are observed similarly to the optical profilometry. Additionally, the aluminium grains can be distinguished, and their observation is enhanced in the Figure III.8, which correspond to a reconstruction of the pad in the aluminium layer from the various SEM views previously taken, following the technique described in the section I.B.I of the second chapter. Large aluminium grains are observed on the pad (on the edges of the reconstruction). In the probe mark, the grains are much finer, which indicates an important work hardening of the aluminium during the EWS. The areas in black on both images correspond to the resin used to perform the observation.



Figure III.7. SEM view of the cross section of the pad after FIB milling.

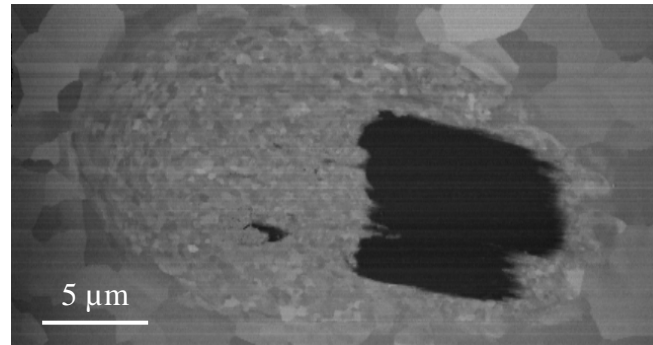


Figure III.8. Reconstructed top view of the pad, in the aluminium layer.

In order to evaluate the plastic deformation undergone by the aluminium during the tests, some EBSD is performed on the pad and in the probed area, using the equipment of the ENSMSE-CMP in Gardanne. Indeed, the plastic deformation is related to crystalline rotation (rotation of 1° ~ elongation of 1%). On the pad surface, the scan is realized on a $10\mu\text{m} \times 10\mu\text{m}$ area to get the initial grain orientation. All the grains have an $\{111\}$ orientation, as depicted by the Figure III. 9. The range of sizes for the aluminium grains using the equivalent circle diameter is found between $0.5\mu\text{m}$ and $3.88\mu\text{m}$.

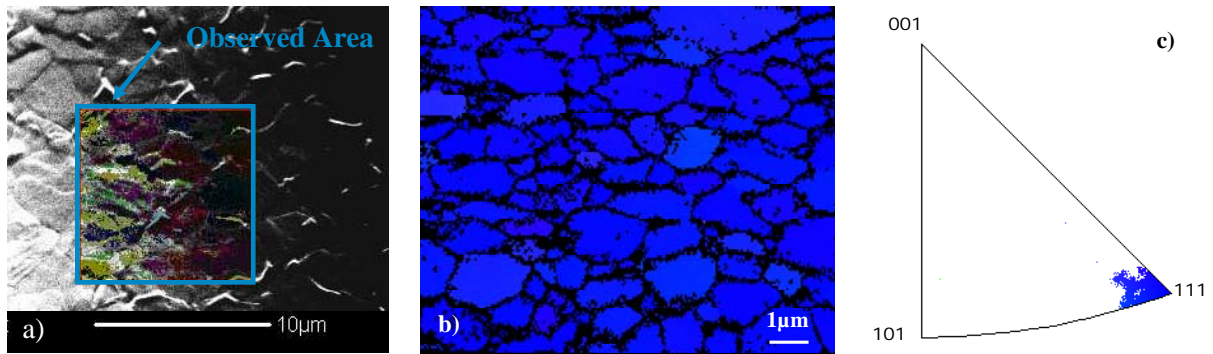
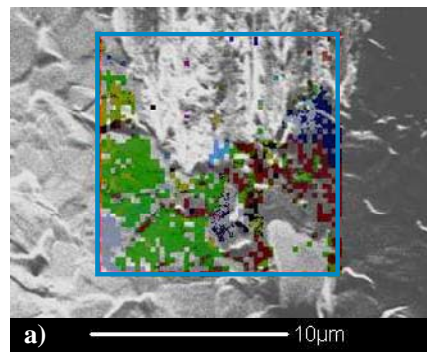


Figure III. 9. SEM top view of the pad (a), grain orientation map (b) and inverse pole figure (c).

Similar analysis is performed in the probed area, delimited by the orange line, and the results are shown in the Figure III.10. However, even after many tests, it seems hard to index most of the points in the probed area. As a matter of fact, the high roughness of the sample, the hardening due to the EWS test and the angle of 70° applied to the sample shadow some part of it. The few indexed points describe the grain rotation, and the new grain orientation. Thus, neither statistics of the grain size nor precise grain reorientation can be deduced. Finally, EWS is shown to induce strong hardening of the aluminium.



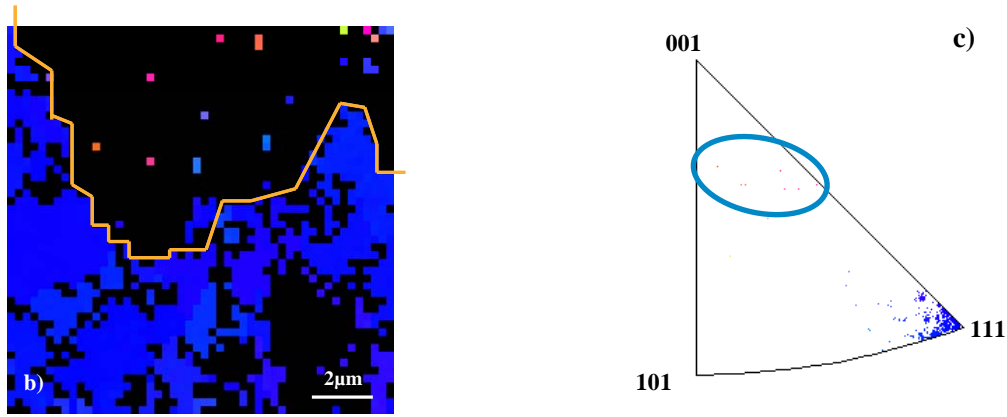


Figure III.10. SEM top view of the probe mark (a), grain orientation map (b) and inverse pole figure (c).

I.B.2. Pad integrity results and observations

The principle of this analysis has been described in the first chapter. In this section, it is performed for various aluminium and padopen thicknesses, which are more specifically considered in the next chapter¹. The thinnest and thickest films are respectively 0.88μm and 1.2μm for aluminium and 0.5μm and 0.7μm for padopen oxide. The analysis is made on 296 pads from two different dies, except for the harshest probing condition whose analysis is made on 592 pads from 4 different dies².

The pad integrity test allows a direct observation of the pad open oxide layer. A SEM observation before and after FIB milling is shown in the Figure III.11. A crack of several micrometres is seen in the oxide layer. This vertical crack goes through its whole thickness.

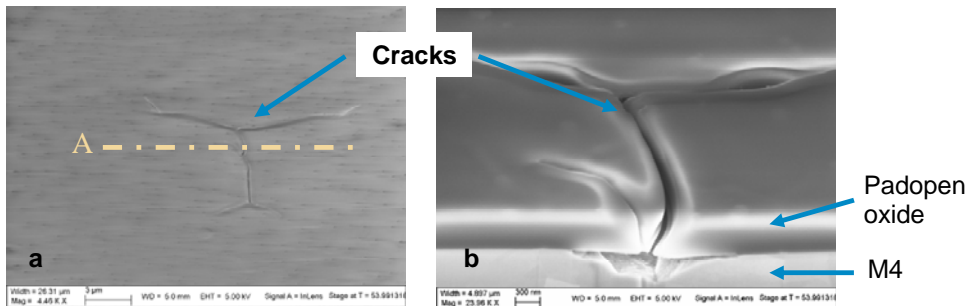


Figure III.11. SEM observations with a 52° tilt of a crack before (a) and after (b) FIB milling along A. The etching of the M4 copper layer (view b) is caused by the chemicals used to ease the crack observation.

	OD 50μm – 5 passes	OD 50μm – 10 passes	OD 90μm – 5 passes	OD 90μm – 10 passes
Alu 1.2 – PO 0.7	0% (0/296)	0% (0/296)	0.7% (2/296)	4.9% (29/592)
Alu 1.2 – PO 0.5	0% (0/296)	0.3% (1/296)	4.1% (12/296)	25.2% (149/592)
Alu 0.88 – PO 0.5	0.7% (2/296)	1.4% (4/296)	7.1% (21/296)	35.8% (205/592)
Alu 0.88 – PO 0.5	0% (0/296)	0% (0/296)	15.9% (47/296)	76.1% (334/444)

Table III.2. Number and rate of cracks observed after pad integrity analysis.

¹ The stack thicknesses are introduced here to further compare the relative effect of the test and the design parameters.

² The enlarged study (OD 90μm and 10 passes) is due to the set up of the plasma etching in the pad integrity analysis for the TaN/Ta removal, for the harshest probing condition and the sample with lowest thicknesses of aluminium and padopen oxide, some pads could not be observed. Hence, only 444 pads have been analyzed on 3 dies (thinnest stack).

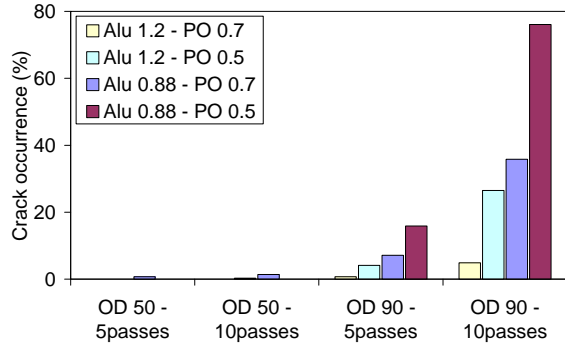


Figure III.12. Crack occurrence for various EWS test conditions.

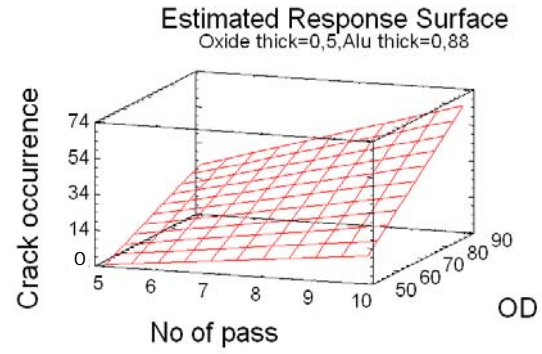


Figure III.13. Estimated response surface for the stack with lowest thicknesses.

The results of the pad integrity analyses are collected in the Table III.2. The increase of the overdrive and the number of pass both lead to an increase of the crack occurrence, as depicted by the Figure III.12. Moreover, the two tests parameters have a cumulative effect as shown by the Figure III.13 and increasing both parameters increases highly the crack occurrence.

I.B.3. Finite element model results and limitations

The effect of the overdrive is evaluated in FE model, keeping the velocity constant. The duration of the application of the velocity depends on the targeted OD. The velocity is applied to the end of the needle. The level 2 needle geometry is used for modelling. In order to evaluate the stress in the pad, the compressive and the Von Mises stresses are investigated. The Von Mises stress or equivalent stress is computed as [Ansys LSDyna, rel. 11]:

$$\sigma_e = \sqrt{\frac{1}{2}[(\sigma_1 - \sigma_2)^2 + (\sigma_2 - \sigma_3)^2 + (\sigma_3 - \sigma_1)^2]} \quad (III.1)$$

where σ_1 , σ_2 and σ_3 are the principal stresses ordered in a way that σ_1 is the most positive (tensile) and σ_3 is the most negative (compressive)

or:

$$\sigma_e = \sqrt{\frac{1}{2}[(\sigma_x - \sigma_y)^2 + (\sigma_y - \sigma_z)^2 + (\sigma_z - \sigma_x)^2 + 6(\sigma_{xy}^2 + \sigma_{xz}^2 + \sigma_{yz}^2)]} \quad (III.2)$$

Since the largest failures are observed in the pad open oxide, the stress analysis focuses on this layer. Hence, the Figure III.14 describes the meshed oxide layer and the elements selected for the analysis. The Figure III.15 shows the compressive stress (in absolute value) in the elements of the top of the padopen layer during one touchdown. Both the compressive and the Von Mises stresses have similar evolutions during the test. During the loading step to reach the maximum overdrive, a given element is stressed when the tip goes over, and when the tip is far enough, the elastic recovery occurs and the stress in the element returns to null. Similar phenomenon is observed during the unloading stage, explaining the two peaks observed. The maximum of the stress undergone by the structure is located in the element where the tip stops at the end of the loading step, and during the plateau.

On that chart, a curve that passes through all the maximum of the successive elements have exactly the same shape than the loading curve and the vertical reaction force. Hence, the compressive and Von Mises stresses in the pad are proportional to the applied load. Since it has been demonstrated in the section III.B.1 of the chapter 2 that increasing the overdrive is related to an increase of the vertical reaction force, the increase of the overdrive leads to an increase of the stress in the structure.

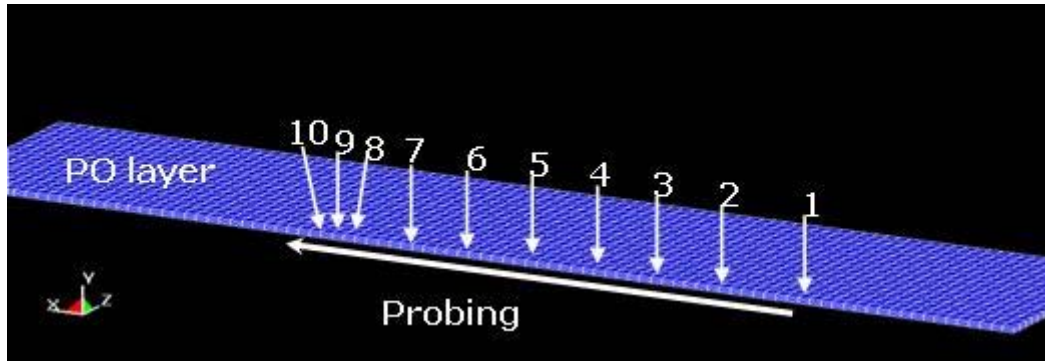


Figure III.14. Meshed padopen oxide layer and elements selected to evaluate the stress in the pad.

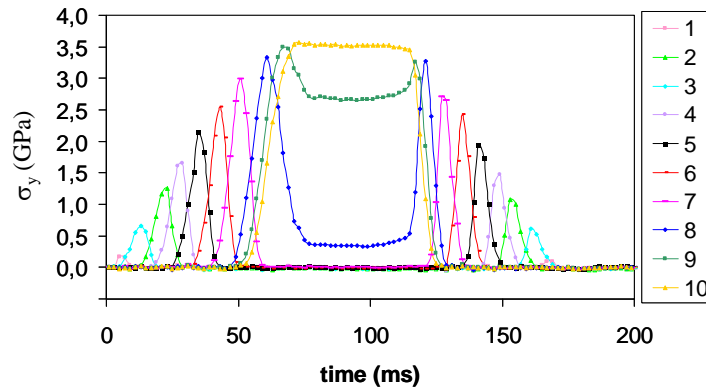


Figure III.15. Compressive stress in the structure during the test.

The number of passes is not estimated with this model for two main reasons. Firstly, elastic material properties are selected, hence at the end of the first touchdown, no stress due to the pass is observed in the structure. Secondly, the current model is quite long to solve (4 days with four processors), and adding plastic material properties to perform a second touchdown plus several passes would take a very long CPU time. Hence this model is not directly able to evaluate the effect of the number of passes on the pad robustness.

However, it has been emphasized in the section III.B.2 of the chapter 2 that increasing the number of passes is related to an increase of both vertical and lateral forces until 5 passes. As a consequence, similarly to the observation for the overdrive, the stress in the structure is expected to rise while increasing the number of passes.

I.B.4. Discussion

In the section I.B.2, the layer stacking options have briefly been shown to be key parameters for the pad robustness. Hence, in this part, the pad design options and the test parameters are weighed to compare their influence and further avoid failures in the structure.

The results are analyzed and trends are extracted with the help of Statgraphics release 5.1. Then, a Pareto chart showing the effect of the considered parameters is depicted in the Figure III.16. Hence the tests parameters and their interactions are shown to be the most important contributors to the appearance of failures in the pad. Moreover, a ranking is established according to the higher ability to induce cracks in the structure to further prevent issues:

Overdrive > Number of passes > Pad stacking options (layer thicknesses)

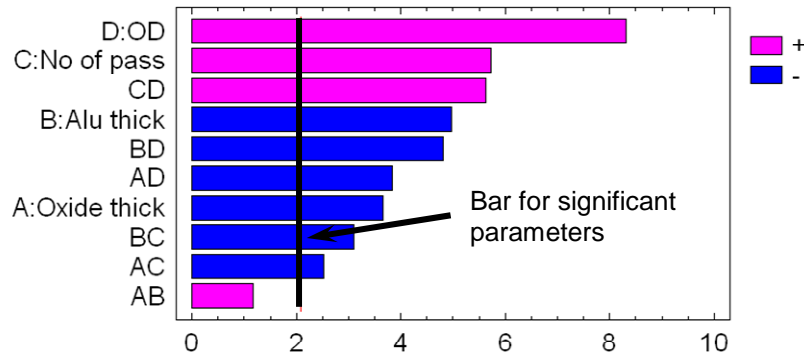


Figure III.16. Standardized Pareto chart for the crack occurrence.

An analysis of variance is made on the results of the design of experiments to check the validity of the model. The coefficient of determination of the model is quite good and r^2 is equal to 0.919.

In the literature, Liu et al. [Liu, 2005; Liu, 2009] performed a probing DOE to weight tests (OD and number of touchdowns) and designs (dielectric thickness and metal thickness) parameters. Their range of parameters is slightly different because the OD varies between 100 μ m and 200 μ m and the number of touchdowns between 1 and 4. Yet, they obtained that the tests parameters and particularly the overdrive are more prone to induce failures in the dielectric layer than the pad layout.

Finally, in both the literature and the current study, the increase of the overdrive and the number of passes is related to a high increase of the fracture hazard. However, these test parameters are often specified by the product's customer¹ or needed to reduce rejects and generally, they cannot be lowered.

II. Gold bonding: Influence of the process parameters

During the wire bonding assembly, several options such as the wire material, the wire diameter and bonding parameters have to be optimized to achieve a good adhesion and electrical connection between the pad and the packaging. Most of the parameters that have been described in the first chapter are considered in this section.

Nowadays, most of the wire connections are made using gold. Hence, bonding with 25 μ m gold wires have been made in a first attempt. This diameter is often employed for CMOS 120nm dies since the mashed ball corresponds to the adequate pad opening. Then, with regards to the results obtained with these wires and the objectives, 20 μ m wires are used. This diameter best suits with more recent technology nodes whose pad openings are smaller. Finally, thermal aging is performed to evaluate the effect of thermal heating on the aging of wires.

II.A. Gold bonding with 25 μ m diameter wires

During the gold bonding study, the Bonding Power (BP), Bonding Time (BT) and Bonding Force (BF) are investigated. The first one corresponds to the ultrasonic power applied to weld the ball to the pad. The second one is the duration of the application of this ultrasonic power. The third one corresponds to the force applied to the ball by the capillary during the bonding process.

¹ Other parameters not discussed here such as the temperature during EWS testing are sometimes additionally specified for particular applications when dies encounter thermal cycles during their lifetime.

II.A.1. First bonding Design of Experiments

During the first wire bonding DOE, bonding is made on probed pads (8 passes at OD 80 μ m) as shown in Figure III.17. The gold bonding is performed with wires manufactured by Kulicke & Soffa. It is a FP2 gold wire (purity: 99%) whose elongation and breaking load are respectively in the ranges 2% to 6% and 13gf to 18gf¹. The bonding is performed using an ASM Eagle 60 bonder at ST Grenoble by the ATM team. The substrate is heated at 180°C during the bonding.

The design of experiment is a full central composite type duplicated two times to get reliable trends and a reasonable amount of wires to pull (about 60 for each configuration), based on the three process parameters. This DOE ensures the isovariance and the orthogonality and is made from fourteen tests plus three repeated for the standard process condition (central point). The standard process conditions for BT, BP and BF are shown in Table III.3.

	Lowest value	Medium value	Highest value
Bonding time (ms)	12	15	18
Bonding power (mW)	46	51	56
Bonding force ² (gf)	17	20	23

Table III.3. Set of bonding parameters for the first DOE.

Once the bonding operation is achieved, the balls heights and diameters are measured and the wires are pulled. For these tests, peeling cases are preferred to neck break ones in order to discriminate the critical bonding parameters. Both failure modes are observed as shown in Figure III.17.

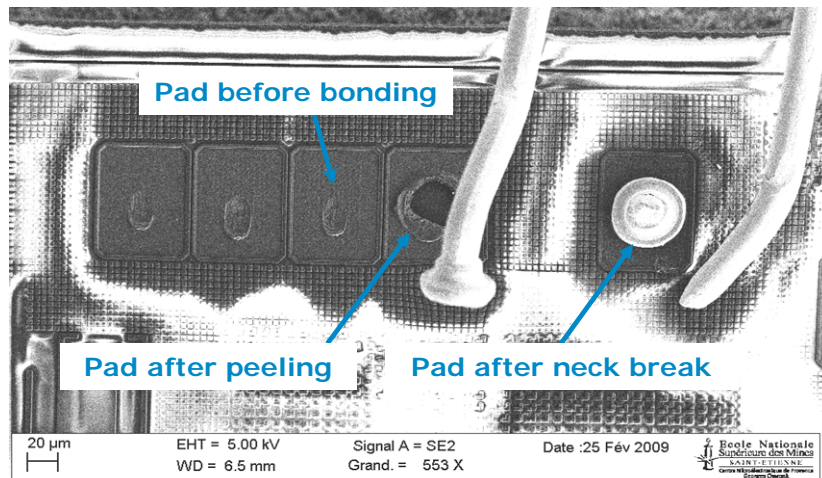


Figure III.17. SEM top view of the pads, showing the probe mark obtained prior to bonding and the 2 failure modes obtained after wire pull test (peeling and neck break).

II.A.1.a) Mashed ball measurements

The measurements of the ball size are made by means of optical profilometry on 3 balls. However, as depicted by the Figure III.18 and Figure III.19, there is no clear trend on both the ball diameter and the ball height while modifying the bonding parameters (BT, BP or BF).

¹ The gram force is commonly used by wire bonding manufacturers and its conversion to the standard unit system is 1gf = 9.8mN.

² In millinewton, the bonding forces are respectively 167mN, 196mN and 225mN.

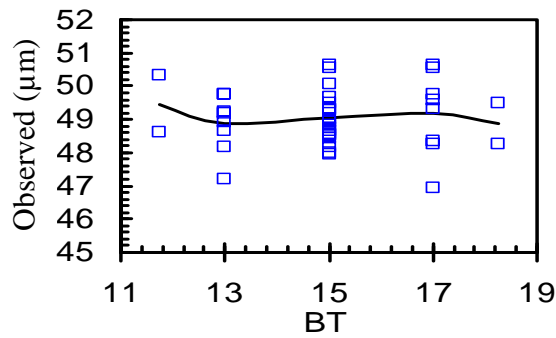


Figure III.18. Variation of the ball diameter as a function of the bonding time.

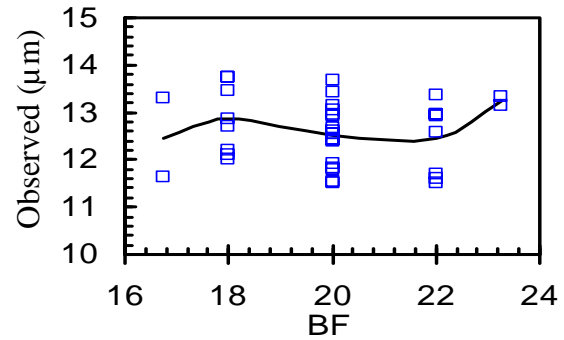


Figure III.19. Variation of the ball height as a function of the ball force.

In the literature, many studies have been led to evaluate the effects of the bonding parameters on the gold ball dimensions. In two successive designs of experiments, Shu et al. [Shu, 1991; Shu, 1996] show that the gold ball diameter in both x and y directions increases while increasing the BP, BF and BT. In their paper, Beng Teck et al. [Beng Teck, 2006] evaluated a new structure for high temperature applications (Electroless Nickel Immersion Gold pad: ENIG pad). Similarly to Shu et al., they found that the gold ball diameter increased while increasing bonding force or bonding power. Kripesh et al. [Kripesh, 2002], demonstrated that increasing either the bonding force or the bonding power is related to a reduction of the ball height. In their study, Gillotti et al. [Gillotti, 2002] showed that an increase in ultrasonic current or BF leads to an increase of the ball diameter. They also found that increasing the BF leads to a decrease of the ball height but an increase of the ultrasonic generator current lead to an increase of the BH. In their paper, Mayer et al. [Mayer, 2003] stated that the increase of the amplitude of the ultrasonic vibrations and the bonding force are both related to a reduction of the ball height and an increase of the ball diameter. Finally, the studies previously described and also many others [Huang, 2009; Breach, 2010] show that the increase of the process parameters is generally related to an increase of the ball diameter and a decrease of the ball height.

Hence, some questions are raised on the range of values for the parameters of the DOE, and also on the ball dimension measurement technique comparing with expected results aforementioned (effects of the bonding force and power on the ball height and diameter).

II.A.1.b) Wire pull tests

During the wire pull tests, 1580 wires are manually pulled using the Dage 5000 described in the first chapter. Among these wires, 96.5% of the failure modes are neck breaks, i.e. fracture in the heat affected zone, and the remaining 3.5% (55 occurrences) are peeling cases. The latter, which is caused by the weakness of the structure, has been more precisely observed in the Figure III.20. Peeling is accompanied by a removal of some of the interconnect structure from the pad, and the border of peeled area raises, showing a delamination within the inner layers of the pad. In order to evaluate the location of the delamination, some optical profilometry is made. As shown in Figure III.21, the profile of the peeled pad underlines several characteristics. Firstly, the top layers surrounding the failures are raised consequently to the failure (about 3 μm). As a consequence, the failure seems to start from the bottom of the pad, in a weak interface, and then reach its surface. Secondly, the delamination is located 4.07 μm under the pad surface, which is linked to a delamination occurring between the M2 and IMD2 layers. Moreover, with respect to the profile at the bottom of the failure, this delamination takes place between the copper of the M2 level and the oxide of the M2 and IMD2 layers.

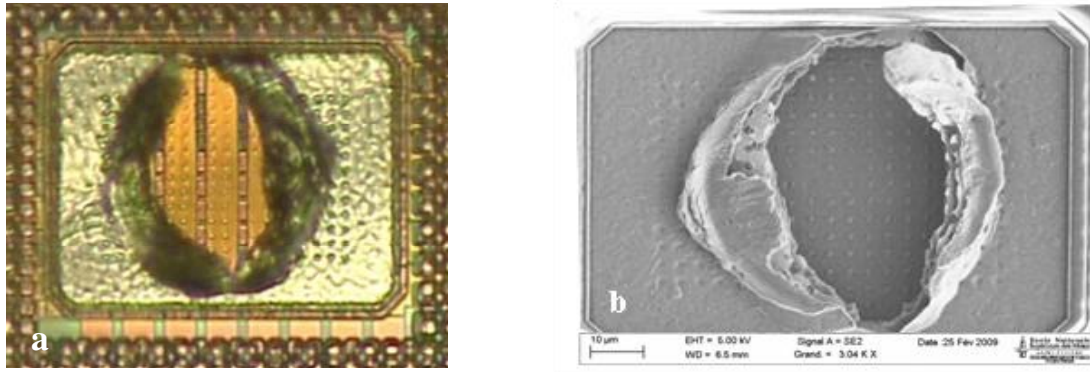


Figure III.20. Top views of pads with peeling: with an optical microscope (a) and with a SEM (b).

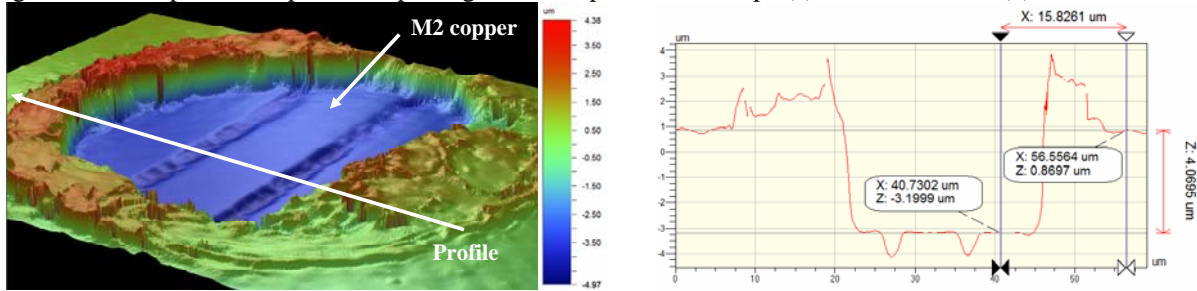


Figure III.21. Optical profilometry image of the peeled pad and profile of the failure.

The peeling occurrence is plotted for all the process conditions, as depicted by the Figure III.22. No clear trend can be found between the peeling occurrence and any of the bonding parameters.

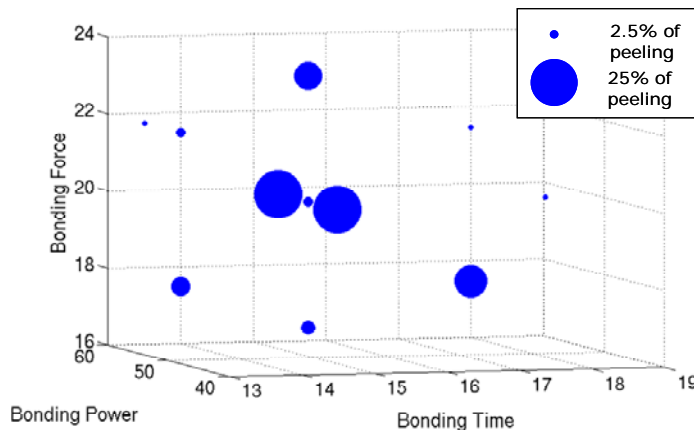


Figure III.22. 3D plot of the peeling occurrence as a function of the bonding parameters.

In the literature, the experimental and numerical studies performed to evaluate the effect of the bonding parameters on the peeling occurrence for gold wire bonding are summarized in the Table III.4. In a general manner, all the works show that the bonding power, time and force have to be reduced to avoid the peeling risks. A purer wire is also advised to reduce the peeling occurrence. Additionally, Gu et al. [Gu, 2003] have shown that the FAB size have to be increased to reduce the fracture hazard. Through a numerical study of the stress in the pad, Liu et al. [Liu, 2004] have found that increasing the ultrasonic amplitude or frequency lead to increase the stress in the pad. Thus, the peeling occurrence is expected to increase as well.

	Variation of the bonding parameters required to reduce the peeling occurrence				
	Impact force	Bonding time	Bonding power	Bonding force	Wire purity
[Gu, 2003]	n/a	Reduce	Reduce	No effect	Increase
[Binner, 2004]	Effect observed: control required	n/a	Reduce	n/a	n/a
[Fiori, 2007]	Increase	n/a	n/a	n/a	Increase
[Zhang, 2007]	n/a	n/a	Reduce	Reduce	n/a

Table III.4. Bonding parameters required variation to reduce the peeling occurrence.

As a consequence, the reliability of the current DOE is questioned since no variation of the ball size and the peeling occurrence are observed while modifying the bonding parameters. Hence, a second, wider, design of experiment is required to investigate bonding parameters, wire material (copper versus gold) and design parameters.

II.A.2. Second wire bonding DOE

In this second DOE, in a first attempt to fix the DOE borders, only the standard and the extreme bonding conditions are made. The same bonder (ASM Eagle) at ST Grenoble and the same gold wire than during the first set of experiments are used. The EFO current for the FAB formation is 3850mA and the EFO duration is 450 μ s. The standard case is also the same. However, the gap with other bonding cases is larger, as shown in Table III.5.

	LLL	MMM	HHH
Bonding time (ms)	13	15	17
Bonding power (mW)	44	51	57
Bonding force ¹ (gf)	15	20	25

Table III.5. Bonding set of parameters for the second design of experiments.

For this second DOE, and all the following gold and copper wire bonding results, the pad organisation on the die is not common (i.e. not a conventional perimeter as defined in the chapter 1). Thus, as shown in the Figure III.23, the wires are long, with minimum, average and maximum sizes of respectively 4466 μ m, 4760 μ m and 5234 μ m. Hence two kinks are made between the pad and the leadframe to achieve a good bonding. The structures are CMOS 120nm structures, made from four copper layers and an aluminium layer on top, as further described (chapter 4 section II.A).

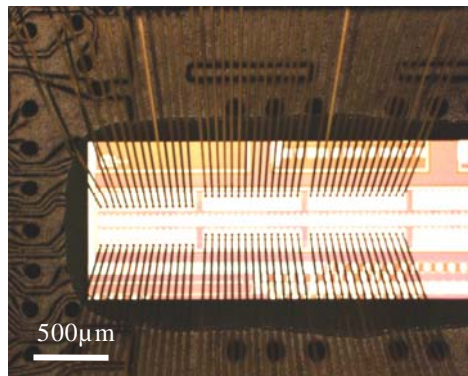


Figure III.23. Optical microscopy top view of the bonded pads.

¹ In millinewton, the bonding forces are respectively 147mN, 196mN and 245mN.

II.A.2.a) Mashed ball measurements

As a high uncertainty has been observed for the measurement technique and the number of measured balls during the first DOE, these are made with an optical microscope on 8 mashed balls. For the ball height, it achieved by successively focussing on the pad and on the collar of the mashed ball.

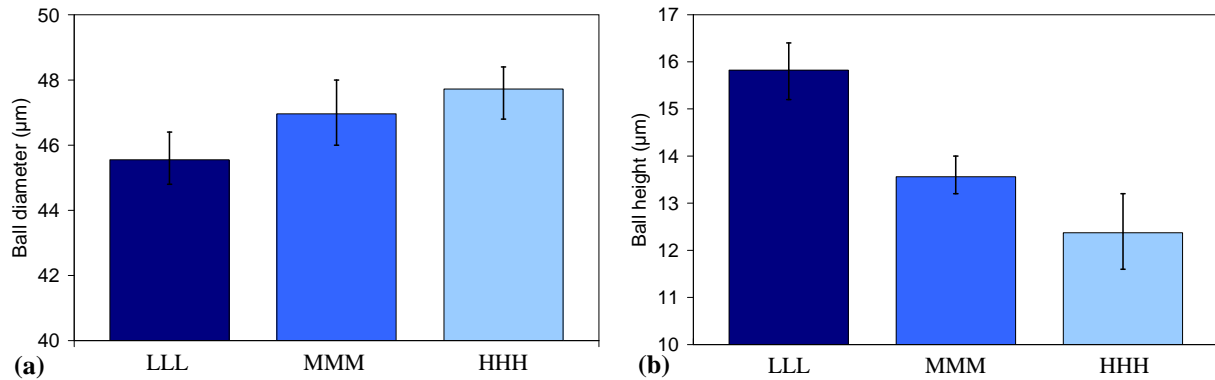


Figure III.24. Effect of the bonding parameters on the ball diameter (a) and the ball height (b).

The measurements of the balls diameters and heights are depicted in the Figure III.24. As expected and found in the literature (see section II.A.1.a)), increasing all the bonding parameters induces an increase of the ball diameter and a decrease of the ball height. Hence, higher confidence is reached for the ball measurement technique and the wire bonding, and the closeness of the values during the first set of parameters is clearly underlined.

II.A.2.b) Results of the wire pull and shear tests

For each set of bonding parameters, 45 wires (15 wires on 3 different dies) are pulled. Only failures at neck break are obtained, whatever the bonding conditions. The range of forces for the neck break cases during the pull tests is between 8.74 gf and 9.95 gf, which are, as expected, below the breaking load of the gold wire.

Additionally, some shear tests are performed to check the bonding resistance. In each case, the ball is removed completely from the pad (failure mode 1) showing average bonding quality of the ball and a reduced formation of IMC. The range of shear forces obtained is between 16.7gf and 23.4gf. The shear strength is calculated from the equation (I.2), and values between 94MPa and 132MPa are obtained.

II.A.3. Discussion

The requirements of the study are not fulfilled with this wire because no peeling is obtained. During the first DOE, some peeling has been found for similar bonding conditions. Hence, two reasons can explain the current results. Firstly, during the first DOE, EWS was made prior to bonding and may have induced pre-cracks in the structure. In the literature, several studies point out the cumulative effect between the probing and the wire bonding. Hotchkiss et al. [Hotchkiss, 2001], showed that increasing the number of touchdowns from 0 (i.e. no EWS) to 4 is directly related to an increase of the number of metal lift failure modes. To prevent issues due this testing and assembly, Yong et al. [Yong, 2003] even proposed a new structure separating the bonding and the probing regions to avoid failure during bonding due to probing. The second reason can be that the current pads are more robust than those of the first DOE. Yet, due to uncertainty on the structures firstly tested, this assumption is hard to check.

As a consequence, for further experiments on gold bonding, smaller wire diameters are used. This modification combines two advantages. The current trend is to move forward to smaller pad pitch, which requires bonding with smaller wires to get smaller mashed balls. Moreover, recurrent issues are obtained while reducing the wire diameter, as shown by Mercado et al. [Mercado, 2000]. The root cause of these issues is that to maintain a stable and robust bond (i.e. shear strength), the wire bond force per unit area is higher for smaller balls than for larger balls. As a consequence, higher stress is undergone by the pad structure and peeling is more likely observed. Thus, for the following wire bonding experiments, a gold wire with a 0.8 mil (or 20 μm) diameter is selected.

II.B. Gold bonding with 20 μm diameter wires

The gold wire for the bonding is still manufactured by Kulicke & Soffa. The wire purity is still 99% (GPG2 wire), and the elongation and breaking load ranges are respectively 4-8% and 5gf (49mN) to 10gf (98mN). Hence, compared to the larger wire, the elongation prior to failure is higher. As expected with the reduction of the wire diameter, the breaking load decreases.

	LLL	MMM	HHH
Bonding time (ms)	16	18	20
Bonding power (mW)	41	45	49
Bonding force ¹ (gf)	6	9	12

Table III.6. Bonding set of parameters for the design of experiment with 0.8 mil gold wires.

The standard case has to be redefined to obtain the desired ball diameter and avoid non sticking on pad cases. Once more, only the 3 extreme bonding set of parameters have been used, as displayed in the Table III.6.

II.B.1. Bonding parameters effects on mashed ball dimension

The effect of the bonding parameters is investigated the same manner as for 25 μm wires. As shown in the Figure III.25 and previously observed for larger wires, increasing altogether the bonding parameters is related to an increase of the ball diameter and a decrease of the ball height. For these smaller wires, the ball diameter and height are smaller as well, of respectively about 12 μm and 4 μm .

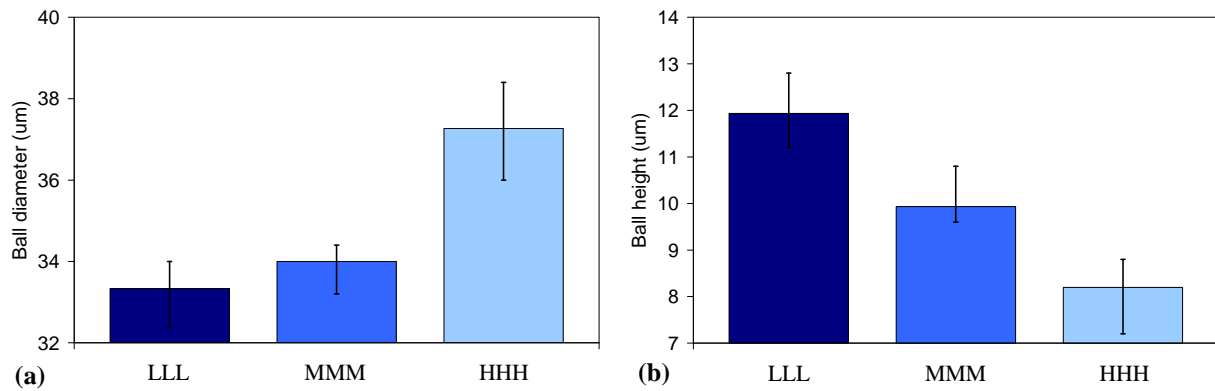


Figure III.25. Effect of the bonding parameters on the ball diameter (a) and the ball height (b).

¹ In millinewton, the bonding forces are respectively 58.8mN, 88.2mN and 117.6mN.

II.B.2. Wire pull tests results

For each set of bonding parameters, 15 pads on 3 different dies are pulled. Only the neck break failure mode is obtained, even for the harshest bonding conditions. The range of forces for the pull tests is very narrow (between 6.04 gf to 6.33 gf), so no trend can be deduced for the effect of the bonding parameters on the pull force. These forces are lower than those of larger wire, which is due to a lower breaking load of the gold wire.

II.B.3. Ball adhesion on pads

The adhesion of the ball to the pad is tested on 180 wires for each set of bonding parameters. Two failure modes are observed during the shear tests. The failure mode 1, depicted in the Figure III.26 (a), is observed on most of the pads. For some tests, gold from the mashed ball remains on the pad after the tests, which corresponds to the failure mode 2, as shown in the Figure III.26 (b). As previously mentioned in the first chapter, the failure mode 2 is preferred since it is related to an enhanced bonding quality, with a noticeable IMC formation and growth during the process.

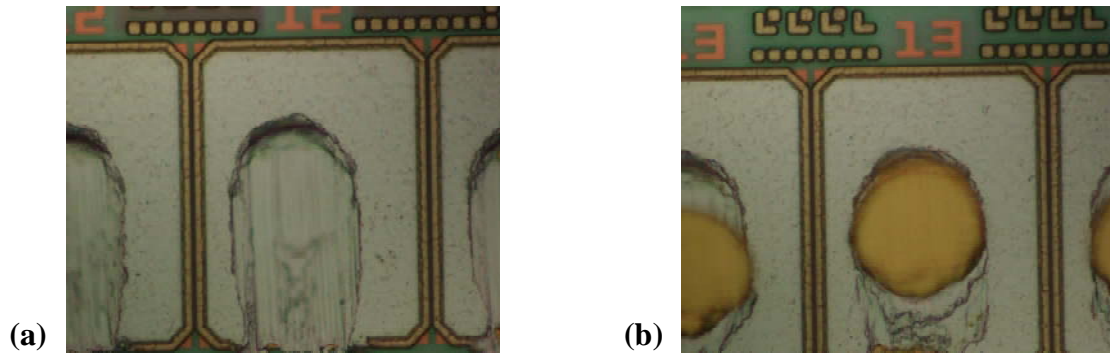


Figure III.26. Failure mode 1 (a) without gold remaining on the pad and failure mode 2 (b) with gold remaining.

As shown by the Figure III.27, increasing the bonding parameters, in the considered range of values, induces an increase of the number of the failure modes 2 observed. Therefore, it enhances the strength of the bond. The bonding time, force and power are coupled in the current study, thus it is not possible to evaluate the significance of the parameters individually.

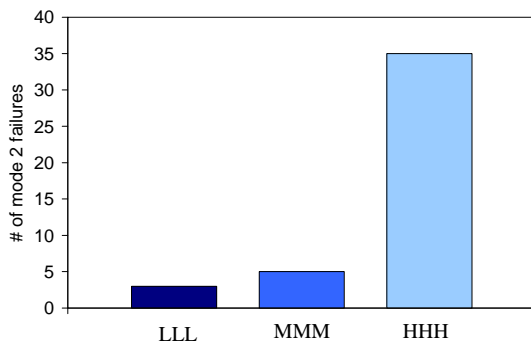


Figure III.27. Number of failure mode 2 for the bonding parameters considered.

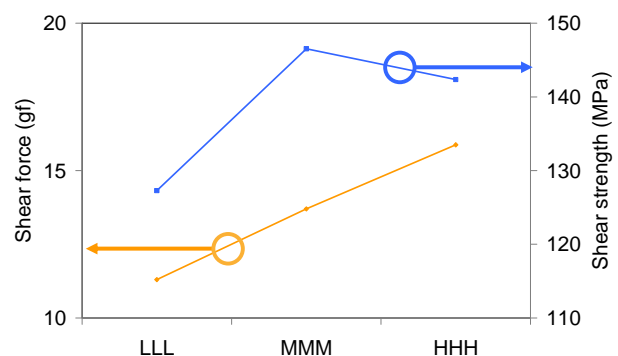


Figure III.28. Average shear stress and shear strength for the set of bonding parameters.

The Figure III.28 shows the average shear force and strength for the set of parameters. The shear strength is calculated using the results of the Figure III.25 (a). The shear force increases with increasing the bonding parameters. The shear strength does not behave similarly. A

maximum is found for the medium bonding parameters. Considering the number of failure modes 2 obtained for the highest set of bonding parameters, this set of parameters is expected to have the highest shear strength. Hence, it seems that the measurements of the balls diameter have once more to be questioned and either the ball diameter of the medium bonding parameter is under evaluated or the one of the high bonding parameters is over evaluated.

In the literature, Zhang et al. [Zhang, 2007] found that the ball shear strength increases when the ultrasonic power increases from lower to medium values, then, while continuing to increase the bonding power, no or little variation is found. Nevertheless, they found no clear correlation between the bonding force and the shear strength. Kripesh et al. [Kripesh, 2002] found that increasing the bond force or the bond power is related to an increase of the shear force, similarly to the results aforementioned.

II.B.4. Discussion

The experimental measurements enable us to observe the effect of the wire diameter on the strength of the bond and the tensile load applied to the pad during the bond testing. Higher pull and shear forces are obtained for the larger wire diameters. Yet, the use of the shear strength as an indicator of the bonding quality seems particularly efficient to evaluate the robustness of the bond since the shear strength of the smaller wire is higher than that of the larger one. This trend is correlated to the failure mode observed for the diameters.

Similarly to the shear strength, while comparing wires with very different diameters (for instance 25 μm wires versus 20 μm wires), it can be interesting to calculate the ratio of the pull force over the application diameter to evaluate the tensile stress applied to the pad during the test. Hence, the equation giving the pull stress is:

$$PS = \frac{4PF}{\pi * D^2} \quad (\text{III. 3})$$

where PS is the pull stress, PF the pull force and D the diameter of the mashed ball, which is assumed to be circular.

The calculation, averaging the ball dimension for each wire diameter, leads to stress ranges from 50MPa to 57MPa and from 59MPa to 62MPa respectively for the large and the small wires. It confirms that higher tensile loading is required to remove the smaller wires. As a consequence higher stress is also undergone by the structure during the testing operations.

Additionally, the IMC coverage would be a good confirmation of the effect of the bonding parameters on the robustness of the bond; however this technique is not very easy to set up. All these measurements raise questions about the validity of wire pull and shear tests to evaluate the bonding quality and the impact of the process on the structure. Indeed, in the literature, Harman et al. [Harman, 1978; Harman, 1997] have shown the parameters and uncertainties related to these tests. Chemical deprocessing can also be an interesting method to evaluate the presence of cracks in the pad structure after wire bonding, avoiding pull and shear tests that induces additional (and different, e.g. tensile loading for the pull tests) stresses into the structure.

II.C. Effect of the thermal aging

The High Thermal Storage (HTS) has been introduced in the first chapter. It is used to age the wire and reproduce the thermal cycles undergone until the end of the chip process and during the die lifetime. The thermal aging implies the IMC growth, which can lead to debonding. However, aging is also related to peeling cases while both the temperature and the time increase [Jamin, 2008]. The samples of the section II.B (20 μm gold wires) are stored during 264 hours at 175°C.

In the literature, Vath et al. [Vath, 2009] found that thermal aging induces an increase of the diameter of the welded ball and a reduction of its height. Nevertheless, for a thermal storage of 264h, the diameter should only slightly increase, independently on the bonding parameters. In the current case, this trend is verified and measurements similar to those depicted in the Figure III.25 are observed.

II.C.1. Wire pull tests results

The wire pull tests are performed on 120 wires for each bonding condition. All but one wire failed at the neck break as observed before aging. In one case shown in Figure III.29, the ball lifted during the test, which highlights the appearance of the voids within the intermetallic compounds.

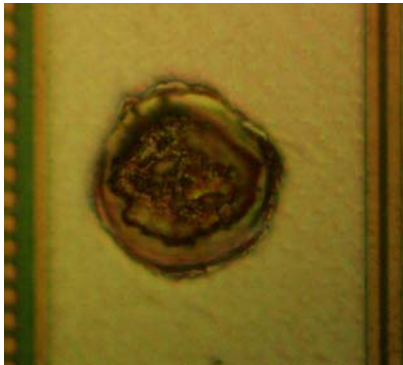


Figure III.29. Optical microscopy top view of a pad after ball lift.

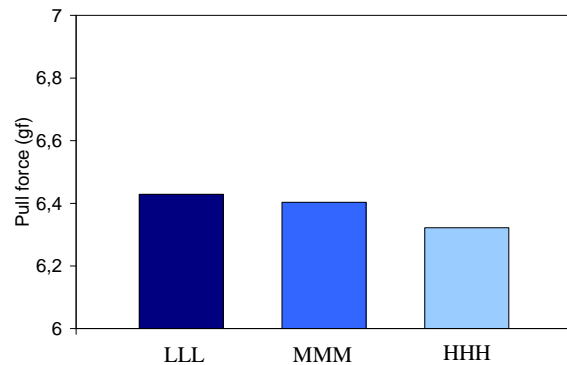


Figure III.30. Measured pull force for the set of parameters investigated.

The Figure III.30 shows that the effect of the increase of the bonding parameters on the measured pull force is almost imperceptible. The average forces are slightly above the ones measured before aging (about 0.2gf or 2mN). In other papers from Jamin et al. [Jamin, 2008] or from Eu et al. [Eu, 2009], the same small increase of the pull force is found while increasing the aging time. Almost similarly, Han et al. [Han, 2007] and Vath et al. [Vath, 2009] found that the pull force is decreasing only a little while increasing the aging time up to one thousand hours. Hence, both current results and published studies lead to quite similar results.

The stability of the pull force during HTS can be explained by the high annealing temperature of the gold wire. In their study, Omar et al. [Omar, 2002] showed that the tensile strength of the gold wire reduces significantly only above 300°C.

II.C.2. Ball adhesion on pads

In parallel to the wire pull tests, 60 gold balls are sheared for the medium set of bonding parameters. The shear forces are in the range 15.4gf to 17 gf, and the shear strengths are between 167MPa and 183MPa. As a consequence, the aging is shown to increase the robustness of the bond due to the growth of the IMC. This corroborates with studies on IMC described in the first chapter.

In their study, Ganesh and co-workers [Ganesh, 2004] demonstrated that increasing the aging time, for a constant temperature of 150°C is related to a small increase of the ball shear force until 2000 hours. On the contrary, Vath et al. [Vath, 2009] showed that the ball shear force firstly increases when the storage time increased, then for more than 500 hours; a sharp decrease of the shear force is obtained, characterizing the ball lifts observed during the pull tests performed in parallel. In a paper, Murali et al. [Murali, 2006] also observed an increase of the shear force while increasing the aging time below 200 hours.

II.D. Discussion

The initial purpose of the second DOE was to enlarge the bonding process window of the first one with other dedicated structures, and to uncouple the EWS and wire bonding effects. The results do not fulfil the expectations of the study since no peeling cases are found. However, the measurements of the mashed ball diameter and height have proven the effect of the bonding parameters, even though only large increase of the bonding power, time and force are required to observe the variations of the ball dimensions. These variations of the wire and the ball diameters highlight the need to use indicators that are independent on the ball size such as the shear and pull forces per unit area. Yet, caution is required on the accuracy of the measurements of the ball size, and as shown in the literature, the use of the IMC coverage instead of the ball diameter to calculate the strengths should improve their accuracy. Moreover, the ball shear tests have underlined the influence of the bonding parameters on the adhesion of the ball to the pad. Due to the current trend to use more and more copper wires to perform the electrical connections, the DOE is extended to copper bonding, which also allows a comparison of the gold and copper materials for bonding.

III. Copper bonding: effect of the process parameters

As mentioned in the first chapter, the actual trend in the semiconductor industry is to move toward copper bonding mainly for cost reduction purposes. Indeed, the increase of the price of the expensive gold becomes a major part of the bonding process cost. Nevertheless, this new bonding material has two major drawbacks, the first one is its increased ball hardness compared to gold. Hence higher bonding parameters are required to form the ball and to realize the bonding. As a consequence, risks of structure damaging are increased. The second one is oxidation in air of the copper. In order to solve this issue, the copper wires used are coated with palladium and/or a forming gas containing N₂ (95%) and H₂ (5%) is used during the bonding operation. These wires are manufactured by Nippon Micrometal Corporation. The wire diameter is kept at 20µm (0.8mil). Its breaking load, elongation and Vickers hardness are respectively 112mN, 11.8% and 55, as given by Uno et al. [Uno, 2009]. Furthermore, they showed the good aging resistance of the copper wire, which is not evaluated hereafter. Similarly, Vath and co-workers [Vath, 2009], compared the IMC growth during thermal aging for gold, copper and palladium coated copper wires. The latter is found to be the most stable and no void is observed either in the wire or at the interface between the pad and the ball.

In a first attempt, the bonding parameters investigated are the same than those of the gold bonding DOE. However, after some initial pull test trials, no peeling cases are observed on the pads. After discussions with the ST Grenoble ATM team, it seems that for copper bonding, another key parameter is presumed to induce peeling during wire pull tests. Hence, a central composite DOE is made to evaluate 2 process parameters (3 levels) and 3 design parameters (2 levels)¹. The contact force varies from 30mN to 50mN and the bonding power from 70mW to 90mW. The bonding time and bonding force are fixed to 12ms and 15mN respectively. The temperature of the strip holder is set to 178°C. The loop height is reduced compared to the gold bonding and fixed at 200µm.

In order to obtain a good reproducibility, for each of the set of process and design parameters, 3 dies of 15 wires are pulled and 2 dies of 15 pads are sheared. As a whole, 3240 wires are pulled and 2160 wires are sheared during the copper bonding design of experiments.

¹ The 3 pad layouts parameters (metal density, presence of vias and pad open vias) are more widely described and studied in the next chapter.

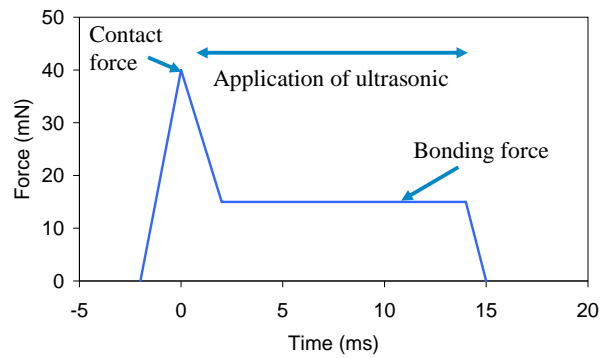


Figure III.31. Schematic representation of the typical forces applied during the copper bonding process.

The Figure III.31 shows the typical forces applied to the pad during the copper bonding operation. The contact force is the highest force applied to the pad, and then the ultrasonic period occurs with the application of a bonding force, and finally, the unloading stage to go to the next bond. The forces undergone by the pad have been measured by Mayer, Shah and co-workers [Mayer, 1998; Mayer, 1998; Mayer, 2003; Shah 2008] for both gold and copper bonding. They used and calibrated microsensors to evaluate the forces during all the bonding stages, and obtained a similar profile. The amplitude of the ball lateral displacement during the application of the ultrasonic power is thus measured between $0.1\mu\text{m}$ and $0.3\mu\text{m}$.

III.A. Initial bonding observations and copper ball measurements

III.A.1. Aluminium pad splash

Once copper bonding is achieved, aluminium from the pad is removed from its original location, squeezed outside the bonded ball perimeter. The aluminium splash results in two hillocks on both sides of the ball, as depicted by the Figure III.32 (a). Since copper is harder than gold (see Figure I.37 in the chapter 1), the aluminium is plastically deformed during the bonding process. According to the SEM view, the splash expands during the application of the ultrasound, and the size of the splash depends on the ultrasound power (more specifically, the amplitude of the ultrasound is suspected), since the splash is observed in the direction of the US. This result is confirmed on the SEM view showing peeling in the structure since the damages of the aluminium are wider in the direction of the ultrasound application, as shown in Figure III.32 (b). The variation of the aluminium splash has been recently studied by Shah et al. [Shah, 2010], who deduced the variations of the splash diameter with the bonding force and the ultrasonic power. They deduced that the major contributor to the splash diameter at interface is the ultrasound applied to the ball, and then in a smaller range the bonding force. Similarly, Clauberg et al. [Clauberg, 2010] found the same dependence of the aluminium splash with the increase of the applied ultrasonic energy.

In a dedicated design of experiments, Qin et al. [Qin, 2009] associated the aluminium splash to a phenomenon of metal wear, and found that the size of the splash depends on the impact force, the bonding force and the bonding power.

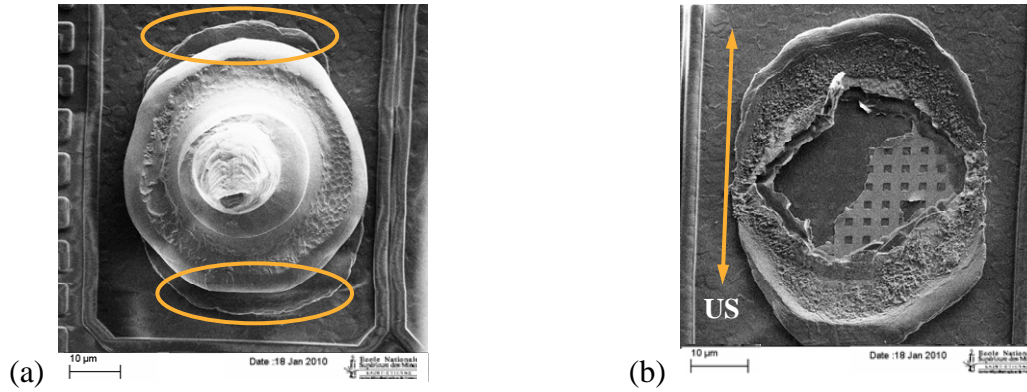


Figure III.32. SEM top view with aluminium splash observed after copper bonding (a) and SEM top view of a peeling case and the direction of the ultrasonic power (b).

Two major issues can arise from high aluminium splashes. Firstly, the thickness of the aluminium layer decreases, which can further induce a mechanical weakness in the structure, and secondly, short circuits can be created depending on the pad pitch and the ultrasonic direction (bridging between the aluminium splashes of two pads).

III.A.2. Mashed ball measurements

The bonded balls are measured with the same tool than the gold balls. Similarly to the first DOE on the gold wires, only little variations are found for the effect of the contact force (CF) and the bonding power (BP) on the ball diameter and height, as shown in Figure III.33. Indeed, the range of sizes for both measurements is very small, almost in the range of the uncertainty of the optical microscope (0.4µm).

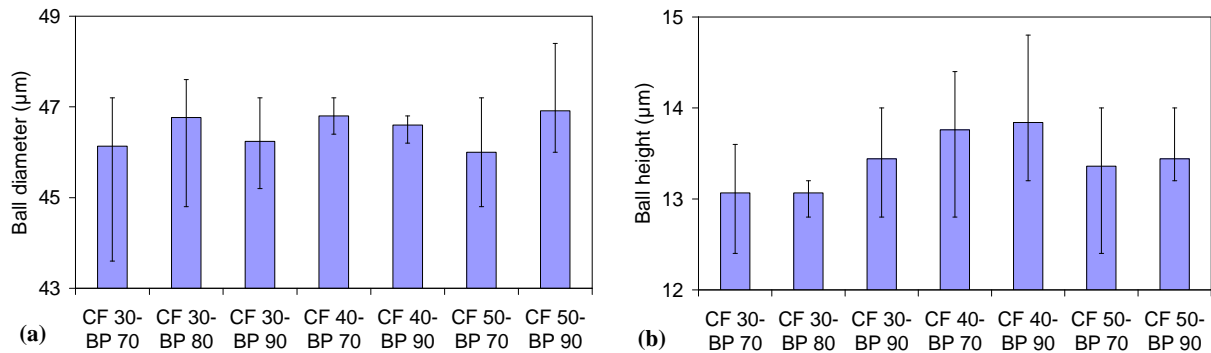


Figure III.33. Effect of the bonding parameters on the ball diameter (a) and the ball height (b).

Similarly to this study, in their work, Shah et al. [Shah, 2010] found little variations of the ball height and the ball diameter over a large range of scale while modifying the ultrasonic power. However, Zhong and co-workers [Zhong, 2007] observed that the major factor on the variations of the balls height and diameter is the ultrasonic power. Its increase is related to a decrease of the ball height and an increase of the ball diameter. In their study, Xu et al. [Xu, 2008] also showed that the ultrasonic power is the most influent factor. Smaller variations of the ball dimensions are observed when the bonding power and the bonding time augment.

III.B. Wire pull tests results

Following the ball measurements, the wires are pulled. Two failure modes are obtained, as expected. On the one hand, the failures at neck break, which occurs in the heat affected zone, similarly to gold bonding, and on the other hand, the failures in the pad structure. The latter failure mode represents about 3% of the wires (98 out of 3240 wires). The Figure III.34 shows

the number of peeling cases observed for all the bonding conditions. Hence, both process parameters are significant. The increase of the contact force and the reduction of the ultrasonic power induce a diminishing of the cracks number in the oxide layer.

The Figure III.35 depicts the safe bonding area to avoid failure risks for a given structure (low metal density, no via and presence of via padopen under the bonding area). Such a 2D contour plot is useful to define a process window to achieve bonding avoiding peeling risks. In the literature, Lim et al. [Lim, 2006] performed a Box-Behnken DOE to evaluate the effect of the bonding time, power and force to set optimized bonding parameters. Similarly, they obtained a safe process window in the centre of the range of values for the bonding parameters.

For the same structure, all the measurements of the pull forces are collected in the Figure III.36, within the ultrasonic power is high to highlight the effect of the contact force. The peeling cases are found for lower pull forces than the failures at neck break. Therefore, increasing the contact force diminishes the peeling occurrence, and also the risks of peeling at reduced pull forces. This trend is highlighted by the box and whisker plot in Figure III.37, within the reduced contact force is related to larger boxes and boundaries. The minimum pull force measured during the tests is 33mN (3.4gf).

In case of failures at the neck, the pull forces are measured between 8gf (78mN) and 9gf (88mN), which is slightly below the breaking load of the copper wire (112mN).

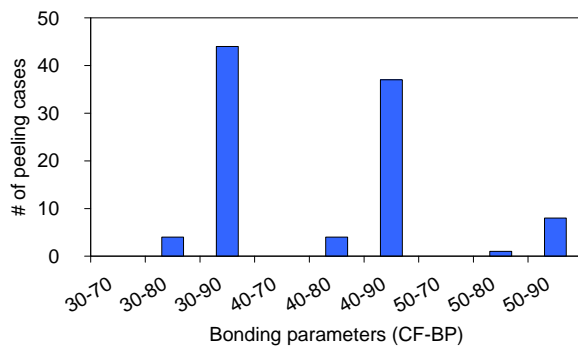


Figure III.34. Number of observed peeling cases for the various bonding parameters.

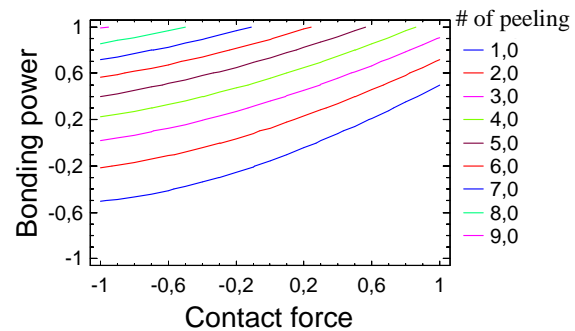


Figure III.35. 2D contour plot of for the effect of the bonding parameters for a given pad structure.

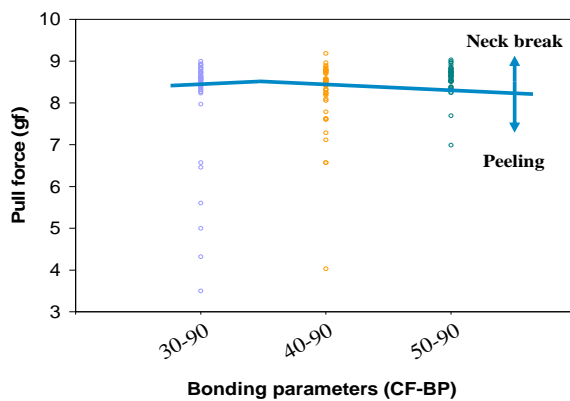


Figure III.36. Measured pull forces for various contact forces.

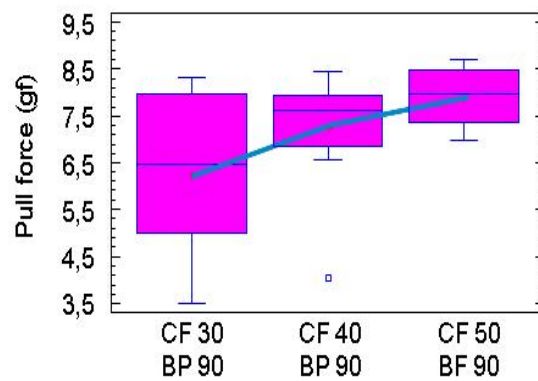


Figure III.37. Box and whisker plot for the effect of the contact force on the peeling occurrence.

During the bonding process, the bonding force is 15gf (or 147mN), the contact force is between 30gf (294mN) and 50gf (490mN). Decreasing the contact force in the range of values considered during the DOE is found to increase the crack occurrence. It is probably due to the formation of the mashed ball during the contact. A higher contact force may induce a larger contact area between the ball and the pad before the application of the ultrasonic power.

Hence, when the ultrasonic power is applied, it is distributed over a larger contact area and finally lower stress is undergone by the pad. In their paper, Qin and co-workers [Qin, 2010] have evaluated bonding parameters with the help of microsensors and laser vibrometer for gold wires. They observed that increasing the contact force induces a larger FAB deformation and as a consequence a larger contact area to end with a similar assumption. They also estimated that the increase of the impact force seems to imply a delay in the bond formation between the pad and the ball. Even if copper does not behave exactly the same way than gold, these observations could confirm the assumptions. One way to check them is to perform copper bonding and only apply the contact force, the bonding time and all the other parameters set to 0. This surely lead to non-sticking on pad cases, but the mark can then be analyzed by means of optical profilometry for instance to check the effect of the impact damages left by the copper ball in the aluminium.

Higher bonding power is applied to the pad and is related to higher shear stresses in the pad structure and then induces the apparition of small cracks or weaknesses in the pad at geometrical singularities that are further highlighted by wire pull tests. The increase of the stress in the structure has been experimentally measured by Shah et al. [Shah, 2009] who found that for bonded balls with equal diameters, the stress in the structure is proportional to the ultrasonic power.

Similarly to this work, Toyozawa et al. [Toyozawa, 1990] established that decreasing the ultrasonic power induces a reduction of the number of cracks in the interconnections. In addition, they found that increasing the bonding load implies an increase of the failure occurrence and when high loads are reached, the fracture hazard diminishes. On the contrary, England et al. [England, 2007] investigated the effects of the bonding power and the bonding force for two strip heater temperatures (150°C and 175°C). They have seen many pad lifts throughout the parameter set, and have not been able to identify key parameters. These peeling cases have been obtained at pull forces close to the ones of failures at the neck break. These are quite higher than the requirements for the robustness of the structure. As a consequence they emphasized that these pad lift failures are not a major concern in such a case.

In a general manner, the peeling cases are critical since they underline a weakness of the structure, however, for quite high pull force values; the pad is robust enough to avoid any failure during its lifetime. Nevertheless, considering also the complexity of the structures, it seems difficult to define a clear threshold that identifies a pad as prone to failures, and generally, peeling failure mode should be avoided.

III.C. Shear tests results

III.C.1. Effect of the shear test parameters on forces

As mentioned in the first chapter, some experimental parameters can influence the measurement of the shear force. As a consequence, the effect of the shear height and the shear velocity are evaluated. The shear height varies from 3 μm to 15 μm . As shown in the Figure III.38, this parameter has little influence on the measurement. For the highest value, a reduction of the force is observed since the shear test is made very close to the top of the ball and does not evaluate the bond anymore.

The shear velocities of 50 $\mu\text{m.s}^{-1}$ and 500 $\mu\text{m.s}^{-1}$ are compared to the standard 250 $\mu\text{m.s}^{-1}$ one. The velocity modifies the results for less than 10%, as shown in Figure III.39. This is mainly observed for low shear velocity, where additional material behaviours are observed due to the slowness of the test. These results corroborate those found by Harman [Harman, 1997].

Hence, the shear height, and the shear velocity do not modify significantly the measurements in the considered range of experimented values, so aforementioned conditions are valid.

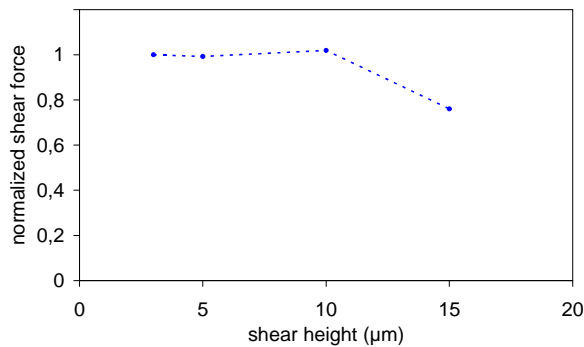


Figure III.38. Variation of the normalized shear force with the shear height.

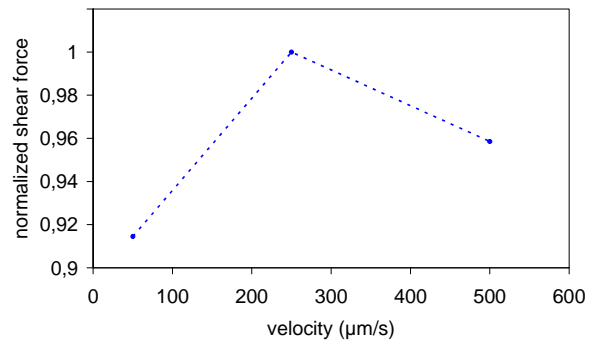


Figure III.39. Variation of the normalized shear force with the shear ram velocity.

III.C.2. Effect of the contact force and the bonding power

The shear tests are performed to evaluate the adhesion of the copper balls to the pads. Only the failure mode 1 is observed (ball lift). The reason for the ball lifts observed is that the copper ball is very strong (for instance compared to gold). As a consequence, the shear ram is not able to shear the ball in its middle, and the copper-aluminium interface is the weakest part of the system, which leads to ball lifts. This failure mode is observed in several studies [Kim, 2003; England, 2007]. Chu et al. [Chu, 2008] explained this failure mode by the low formation of the IMC during copper bonding process.

Increasing the contact force is shown to reduce the shear force, and augmenting the bonding power increases it, as shown by the Figure III.40. Moreover, these two parameters have a cumulative effect, and the highest shear force is obtained while the former parameter is minimized and the latter maximized.

Since no clear trend has been observed for the measurements of the ball diameters, the shear strength is not depicted, however, it is expected that the effect of the contact force is reduced and on the contrary, the effect of the ultrasonic power increased¹. As a consequence, the results of the shear test clearly show an effect of the bonding parameters on the adhesion of the ball to the pad. Considering a ball diameter of 46 μm , the shear strength is found between 136MPa and 171MPa.

It appears that the results of the shear tests are in opposition with those of the pull tests. For gold bonding, usually, ensuring a good ball adhesion is the best way to achieve a good bonding and good formation of the intermetallic compounds. It seems that for copper bonding, due to the material properties of the wire and the bonding parameters, this objective is not valid anymore, and highest shear forces are obtained for conditions inducing the highest peeling occurrence. Similar observation has been obtained by Shah et al. [Shah, 2009].

¹ Considering the results shown by Qin et al. [Qin, 2009], i.e. to maintain the diameter of the copper ball constant, if the contact force decreases, the ultrasonic power has to be increased.

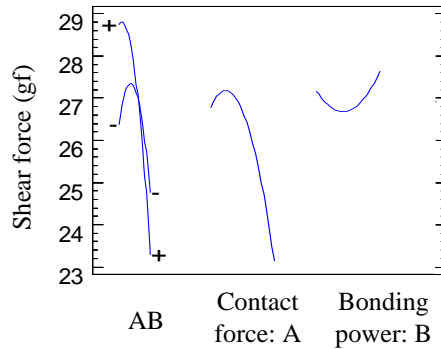


Figure III.40. Effect of the bonding parameters on the measured shear forces.

Some of the recent studies that investigated the effect of the bonding parameters on the value of the shear force and/or shear strength for copper wires are summarized in the Table III. 7. In accordance with the aforementioned work, the increase of the ultrasonic power and the decrease of the contact force enhance the adhesion of the copper ball to the bond pad.

Investigated parameters	Results and observations
Lim et al. [Lim, 2006]	
<ul style="list-style-type: none"> - Bonding power - Bonding force 	<p>The SF is shown to increase while increasing the bonding power. The behaviour for the force is more complicated.</p>
Zhong et al. [Zhong, 2007]	
<ul style="list-style-type: none"> - Bonding power 	Both shear force and shear strength are found to increase while ultrasonic power is increased.
England et al. [England, 2007]	
<ul style="list-style-type: none"> - Bonding power - Bonding force 	<p>SF is found to increase while the investigated parameters augment.</p>
Xu et al. [Xu, 2008]	
<ul style="list-style-type: none"> - Bonding power - Bonding force - Bonding time - Substrate heat 	<p>SF and SS are investigated. Generally, similar trends are observed for both. Increasing the BP or BF is related to increasing and then decreasing the shear values. Increasing the temperature induces an increase as well. Increasing the BT is linked to an increase and then a plateau.</p>

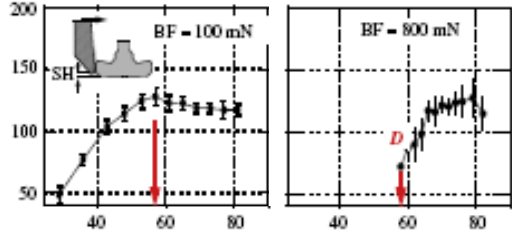
Qin et al. [Qin, 2009]	
<ul style="list-style-type: none"> - Bonding power - Contact force - Bonding force - EFO firing time 	Many coupled parameters: SS increases while: the contact force is reduced and the BP increased, the BF is reduced and the BP increased, the EFO firing time decreases.
Shah et al. [Shah, 2010]	
<ul style="list-style-type: none"> - Bonding power - Bonding force 	<p>SS increases while BP increases and while BF decreases (in a smaller range).</p> 

Table III. 7. Bonding parameters effect on shear force (SF) and strength (SS) found in the literature.

III.D. Discussion

III.D.1. Comparison between gold and copper bonding

During all the study, gold and copper bonding have been successively performed to realize the electrical connections between the die and the substrate. For a similar wire diameter, the copper bonding is found to be able to induce peeling much easily than gold bonding, due to the necessity to increase the bonding parameters to achieve the bonding. The Table III.8 highlights the needs to increase the bonding force and bonding power to achieve the bonding operation while using copper wires. Additionally to these parameters, Shah et al. [Shah, 2008] showed that the impact force has to be increased for copper bonding as well. In the same manner, Qin et al. [Qin, 2009], using the same bonding parameters for both wires, have obtained some peeling case only for the copper wires, similarly to the current study.

	Gold bonding	Copper bonding
Bonding time (ms)	18	12
Bonding power (mW)	45	80
Bonding force (gf)	9	15

Table III.8. Standard bonding parameters used to achieve bonding with gold and copper wires.

The mechanical properties of the gold and copper wires are clearly observed during the pull and shear tests. During the pull tests, the higher elongation prior to failure of the copper is observed during tests whose failures are at the neck. Its higher hardness modifies the failure mode observed during the shear tests.

	Gold bonding	Copper bonding
Wire pull tests failure mode	Neck break	Neck break and peeling
Wire pull force range for neck break (gf)	6.2 to 6.6	8 to 9
Pull strength (MPa)	59 to 62	49 to 53
Shear failure mode	Ball lift and ball shear	Ball lift
Shear force (gf)	15.4 to 17	23 to 29
Shear strength (MPa)	167 to 183	136 to 171

Table III.9. Measurement and failure modes obtained during pull and shear tests for gold and copper bonding.

The differences between gold and copper wires are summarized in the Table III.9. Yet, in this particular case, the shear strength of copper bonds is quite low. As noticed previously, the process window has been changed a little to obtain peeling cases, which can explain the low values obtained. In the literature, many studies reports higher adhesion of the copper ball to the pad [Shah, 2008; Vath, 2009] and found that the copper bonding improves the pull and shear tests performances even though the process window is much smaller to avoid pad lift [England, 2007].

III.D.2. Weight of the process and design parameters

In this section the process and design parameters are faced with respect to wire pull tests (particularly the peeling cases) and shear tests. In that frame, the Figure III.41 show the standardized Pareto charts of the peeling occurrence observed during the copper bonding. The effects of the bonding parameters have been previously described. The bonding power is the major contributor to the appearance of pad lift. The contact force is also pointed out. Some interactions between the process and design parameters are also observed, but they are more widely treated in the next chapter since the pad layout is clearly involved.

The Figure III.42 depicts a similar trend for the shear force. The contact force and the ultrasonic power and their interaction are the preponderant parameters for the adhesion of the copper ball to the aluminium pad. Hence, from theses Pareto charts, the process parameters during wire bonding are always shown to be very critical, and often more than the design ones.

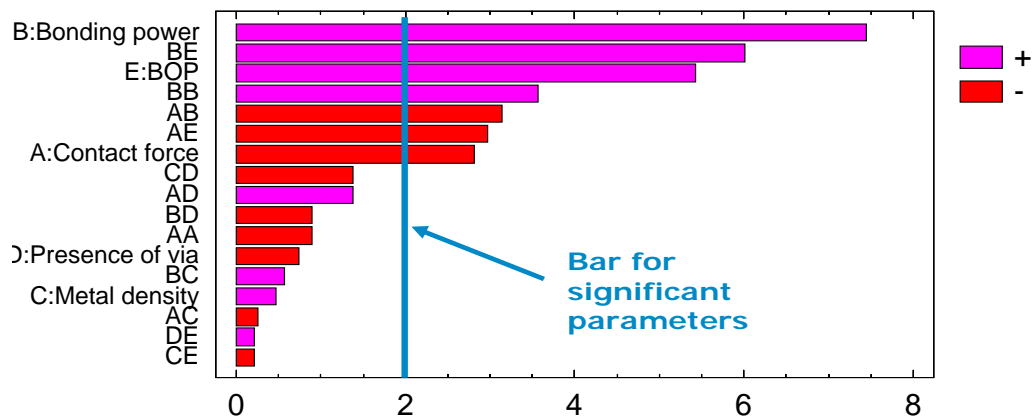


Figure III.41. Standardized Pareto chart of the peeling occurrence after pull tests.

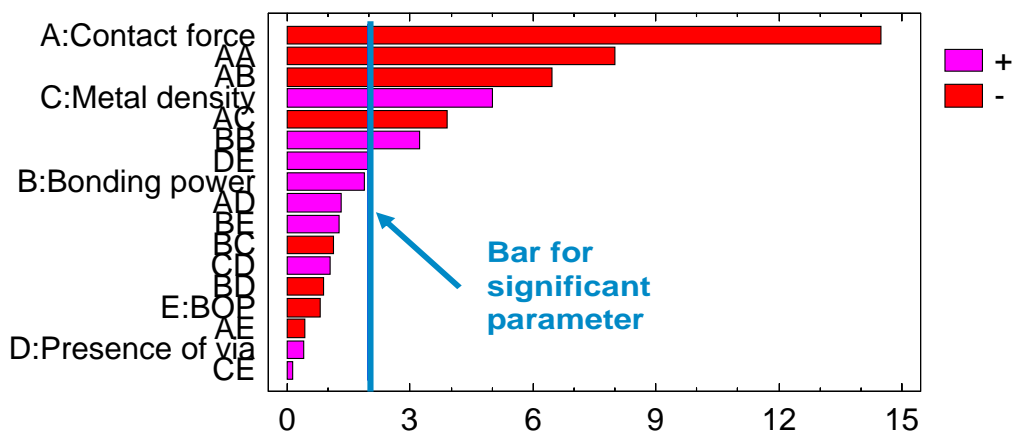


Figure III.42. Standardized Pareto chart of the shear force after shear tests.

IV. Conclusions

In this third chapter, the process and test parameters have been investigated. Firstly, the dynamical effects during the probing have been evaluated. According to the analyses on the reaction forces and the stress obtained from the simulations, the velocity of the chuck and its oscillations during the process do not increase significantly the stress in the pad. As a consequence, they are unlikely responsible for the cracks observed inside the structure. Then, pad integrity analyses on dies and the numerical model show that the increase of the overdrive and the number of passes increases the damages on the aluminium capping layer and also the failure hazard in the pad beneath. Nevertheless, due to the use of elastic material properties, the numerical model is not able to investigate directly the effect of the number of passes.

In a second part, several wire bonding process parameters are evaluated through designs of experiments. The main objective of the studies is to observe peeling during pull tests to find out the parameters detrimental to the pad. For the gold bonding, the bonding time, the bonding force and the bonding power are studied. The ball dimensions depend on the process parameters. Therefore, the increase of the parameters altogether enables us to see an increase of the ball diameter and a reduction of the ball height. The pulls tests only give failures at the wire neck, which do not match the expectations. Then, the shear tests provide some insight of the effect of the bonding parameters on the adhesion of the gold ball to the pad.

Due to the current trend to move toward copper bonding, similar study is performed. To that purpose, the contact force and the bonding power are considered. The small variation of the parameters and the accuracy of the measurement tool have not allowed to establish correlations between the bonding parameters and the size of the ball even if they are expected to evolve in the same way than for gold bonding. Many peeling cases have been obtained during the wire pull tests, which highlight that increasing the bonding power and diminishing the contact force reduces the fracture hazard in the structure. Then, the shear tests showed that contrary to gold bonding, the increase of the bond strength between the copper and the aluminium of the pad is often related to an increase of the peeling risks. Therefore, in this chapter, guidelines for process and test parameters to reduce the fracture hazard in the pad structure are obtained. The necessity to find a balance between the bonding parameters to achieve a good and non damaging bonding is thus underlined.

Finally, it is demonstrated that the process and test parameters show even higher preponderance than the layout ones to imply failures in the pad. However, during wire bonding or EWS, it is generally not possible to modify the parameters aforementioned. Hence, the study of the structure beneath the pad is also very important to avoid failures in the structure and to find optimized structures with enhanced mechanical robustness.

References

[Ansys LSDyna, rel. 11] Ansys LSDyna user manual, release 11.

[Beng Teck, 2006] Beng Teck, Ng, Ganesh, V.P. and Lee, C., Optimization of gold wire bonding on electroless nickel immersion gold for high temperature applications, Electronics packaging technology conference 2006, pp. 277-282.

[Binner, 2004] Binner, R., Schopper, A. and Castaneda, J., Gold wire bonding on low-k material: a new challenge for interconnection technology, IEEE/SEMI International electronics manufacturing technology symposium, pp. 13-17.

[Breach, 2010] Breach, C.D. and Wulff, F.W., A brief review of selected aspects of the materials science of ball bonding, Microelectronics reliability, vol. 50, 2010, pp. 1-20.

[Cacho, 2007] Cacho, F., Vettori, R., Galbiati, G., Thiery, G., Fiori, V., Ainaoui, M., Delpech, P., Tavernier, C. and Jaouen, H., Study of probing test for electrical wafer sort application: dynamical simulation and experimental measurement, IEEE 9th Electronics packaging technology conference, 2007, pp.391-395.

[Chang, 2010] Chang, H.-Y., Pan, W.-F., Shih, M.K. and Lai, Y.S., Geometric design for ultra long needle probe card for digital light processing wafer testing, Microelectronics reliability, vol. 50, 2010, pp. 556-563.

[Chu, 2008] Chu, H., Hu, J., Jin, L. and Jie, Y., Defect analysis of copper ball bonding, International conference on electronic packaging technology & high density packaging, 2008, pp. 1-3.

[Clauberg, 2010] Clauberg, H., Backus, P. and Chylak, B., Nickel-palladium bond pads for copper wire bonding, Microelectronics reliability, 2010.

[England, 2007] England, L. and Jiang, T., Reliability of cu wire bonding to al metallization, IEEE electronics components and technology conference, 2007, pp. 1604-1613.

[Eu, 2009] Eu, P.L., Yong, C.C, Chin, T.S., Tsuriya, M., Nick, V., Lee, B.S., Zul-Kifli Mohd, F. and Fadzli, S., 2N Au wire bonding for ultra fine pitch BGA, IEEE Electronics packaging technology conference, 2009, pp. 486-493.

[Fiori, 2007] Fiori, V., Beng, L.T., Downey, S., Gallois-Garreignot, S. and Orain, S., Gold wire bonding induced peeling in cu/low-k interconnects: 3D simulations and correlations, IEEE International conference on thermal, mechanical and multi-physics simulation experiments in microelectronics and micro-systems, Eurosime, 2007, pp. 1-9.

[Ganesh, 2004] Ganesh, V.P., Kripesh, V., Chih-Hang, T. and Lim, L.A., Wire bonding on a novel immersion gold capped copper metallised integrated circuit, IEEE electronics components and technology conference, 2004, pp. 358-364.

[Gillotti, 2002] Gillotti, G. and Cathcart, R., Optimizing wire bonding processes for maximum factory portability, International electronics manufacturing technology symposium, 2002, pp. 329-334.

[Gu, 2003] Gu, X., Antol, J, Yao, Y.F. and Chua, K.H., A reliable wire bonding on 130nm Cu/low-k device, IEEE Electronics packaging technology conference, 2003, pp. 707-711.

[Han, 2007] Han, M.C., Yan, B.Y., Yao, J.Z., Tran, T.A., Lee, S. and Li, J., Low-k CMOS65 ball grid array 47µm pitch wire bonding process development, IEEE Electronics packaging technology conference, 2007, pp. 613-617.

[Harman, 1978] Harman, G.G. and Cannon, C.A., The microelectronic wire bond pull test – how to use it, how to abuse it, IEEE transactions on components, hybrids, and manufacturing technology, vol. CHMT-1, no. 3, September 1978, pp. 203-210.

[Harman, 1997] Harman, G.G., Wire bonding in microelectronics: materials, processes, reliability and yield, second edition, McGraw-Hill, p. 86.

[Hotchkiss, 2001] Hotchkiss, G., Ryan, G., Subido, W., Broz, J., Mitchell, S., Rincon, R., Rolda, R., and Guimbaolibot, L., Effects of probe damage on wire bond integrity, IEEE Electronics components and technology conference, 2001, pp. 1175-1180.

[Huang, 2009] Huang, W., Computational modeling and optimization for wire bonding process on cu/low-k wafers, International conference on electronic packaging technology and high density packaging (ICEPT-HDP), 2009, pp. 344-352.

[Jamin, 2008] Jamin, L., Atzmon, Z., Dominik, S. and Murali, S., Wire bond reliability – an overview on the mechanism of formation/growth of intermetallics, SEMICON Singapore, 2008, pp. 1-9.

[Kim, 2003] Kim, H.-J., Lee, J.Y., Paik, K.-W., Koh, K.-W., Won, J., Choe, S., Lee, J., Moon, J.-T. and Park, Y.-J., Effect of cu/al intermetallic compound (imc) on copper wire and aluminium pad bondability, IEEE transactions on components and packaging technologies, vol. 26, no. 2, June 2003, pp. 367-374.

[Kripesh, 2002] Kripesh, V., Sivakumar, M., Lim, L.A., Kumar, R. and Iyer, M.K., Wire bonding process impact on low-k dielectric material in damascene copper integrated circuits, Electronic components and technology conference, 2002, pp. 873-880.

[Lim, 2006] Lim, C.C., Ng, K.C., Lee, C.C., Chai, M.K., Lim, O.S. and Chua, K.Y., Copper wire reliability and bonding integrity robustness on cratering sensitive bond pad structure, International conference on electronics manufacturing technology, 2006, pp. 354-364.

[Liu, 2006] Liu, D.S. and Shih, M.K., Experimental method and FE simulation model for evaluation of wafer probing parameters, Microelectronics journal, vol. 37, 2006, pp. 871-883.

[Liu, 2004] Liu, Y., Irving and Luk, T., Thermosonic wire bonding process simulation and bond pad over active stress analysis, IEEE Electronic components and technology conference, 2004, pp. 383-391.

[Liu, 2005] Liu, Y., Desbiens, D., Irving, S., Luk, T., Edborg, S., Hahn, D. and Park, S., Probe test failure analysis of bond pad over active structure by modeling and experiments, IEEE electronic components and technology conference, 2005, pp. 861-866.

[Liu, 2009] Liu, Y., Luk, T. and Irving, S., parameters modeling for wafer probe test, IEEE transactions on electronics packaging manufacturing, vol. 32, no. 2, April 2009, pp. 81-88.

[Mayer, 1998] Mayer, M., Paul, O. and Baltes, H., In-situ measurement of stress and temperature under bonding pads during wire bonding using integrated microsensor, International conference on emerging microelectronic and interconnection technologies, EMIT, 1998, pp. 129-133.

[Mayer, 1998] Mayer, M., Paul, O., Bolliger, D. and Baltes, H., In-situ calibration of wire bonder ultrasonic system using integrated microsensor, IEEE electronic packaging technology conference, 1998, pp. 219-223.

[Mayer, 2003] Mayer, M. and Schwizer, J., Thermosonic ball bonding based on ultrasonic friction power, IEEE Electronics packaging technology conference, 2003, pp. 738-743.

- [**Mercado, 2000**] Mercado, L., Radke, R., Rutson, M., Tran, T.A., Williams, B., Yong, L., Chen, A., and Chen, S., Reliability of multi-layer aluminium capped copper interconnect structures, IEEE International electronics manufacturing symposium, 2000, pp. 84-93.
- [**Murali, 2006**] Murali, S., Srikanth, N. and Vath III, C.J., Effect of wire diameter on the thermosonic bond reliability, Microelectronics reliability 46, 2006, pp. 467-475.
- [**Omar, 2002**] Omar, G., Muhamad, M.R. and Daud, A., Effect of annealing and alloying to the microstructural property of gold wire in wire bond interconnect, Proceeding of ICSE 2002, IEEE international conference on semiconductor electronics, ICSE, 2002, pp. 480-484.
- [**Qin, 2009**] Qin, I., Shah, H., Huynh, C., Meyer, M., Mayer, M. and Zhou, Y., Effects of process parameters on pad damage during au and cu ball bonding processes, IEEE electronics packaging technology conference, 2009, pp. 573-578.
- [**Qin, 2010**] Qin, I., Shah, A., Huynh, C., DeAngelis, D., Meyer, M., Mayer, M. and Zhou, Y., Thermosonic Au ball bonding process investigated using microsensor and laser vibrometer, IEEE electronic components and technology conference, 2010, pp. 1776-1782.
- [**Schmadlak, 2007**] Schmadlak, I. and Hauck, T., Simulation of wafer probing process considering probe needle dynamics, IEEE International conference on thermal, mechanical and multi-physics simulation experiments in microelectronics and micro-systems, Eurosime, 2007, pp. 1-5.
- [**Shah, 2008**] Shah, A., Mayer, M., Zhou, Y., Hong, S.J. and Moon, J.T., In situ ultrasonic force signals during low-temperature thermosonic copper wire bonding, Microelectronics engineering, vol. 85, issue 9, pp. 1851-1857.
- [**Shah, 2010**] Shah, A., Rezvani, A., Mayer, M., Zhou, Y., Persic, J. and Moon, J.T., Reduction of ultrasonic pad stress and aluminium splash in copper ball bonding , Microelectronics reliability, 2010.
- [**Shu, 1991**] Shu, W.K., Fine pitch wire bonding development using a new multipurpose, multipad pitch test die, Electronic components and technology conference, 1991, pp. 511-518.
- [**Shu, 1996**] Shu, W.K., PBGA wire bonding development, Electronic components and technology conference, 1996, pp. 219-225.
- [**Toyozawa, 1990**] Toyozawa, K., Fujita, K., Minade, S. and Maeda, T., Development of copper wire bonding application technology, IEEE transactions on components, hybrids, and manufacturing technology, vol. 13, no. 4, December 1990, pp. 667-672.
- [**Uno, 2009**] Uno, T., Terashima, S. and Yamada, T., Surface enhanced copper bonding wire for LSI, IEEE electronics components and technology conference, 2009, pp. 1486-1495.
- [**Vath, 2009**] Vath III, C.J., Gunasekaran, M. and Malliah, R., Factors affecting the long term stability of Cu/Au bonds subjected to standard and extended HTS, IEEE Electronics packaging technology conference, 2009, pp. 374-380.
- [**Xu, 2008**] Xu, H., Liu, C., Silberschmidt, V.V. and Wang, H., Effect of process parameters on bondability in thermosonic copper ball bonding, IEEE Electronic components and technology conference, 2008, pp. 1424-1430.
- [**Yong, 2003**] Yong, L., Tran, T.A., Lee, S., Williams, B. and Ross, J., Novel method of separating probe and wire bond regions without increasing the die size, IEEE Electronic components and technology conference, 2003, pp. 1323-1329.

[Zhang, 2007] Zhang, L., Gumaste, V., Poddar, A., Nguyen, L. and Schulze, G., Analytical and experimental characterization of bonding over active circuitry, Journal of electronic packaging, vol. 129, December 2007, pp. 391-399.

[Zhong, 2007] Zhong, Z.W., Ho, H.M., Tan, Y.C., Tan, W.C., Goh, H.M., Toh, B.H. and Tan, J., Study of factors affecting the hardness of ball bonds in copper wire bonding, Microelectronics engineering, vol. 84, 2007, pp. 368-374.

Chapter 4: Layout effects on the mechanical robustness of the pads

Outlines

I.	Stacking influence: thickness of the top layers	116 -
I.A.	Failure analyses after EWS	117 -
I.A.1.	Crack occurrences after pad integrity tests	117 -
I.A.2.	3D reconstruction of the cracks under the pads	118 -
I.B.	Resistance to compression: nanoindentation tests	120 -
I.B.1.	Cube corner tip	120 -
I.B.2.	Berkovich tip	125 -
I.C.	EWS numerical model: stress analysis.....	129 -
I.D.	Discussion and correlation between techniques about the thickness effect	130 -
I.D.1.	Comparison between techniques to evaluate the thickness effect.....	130 -
I.D.2.	Discussion	131 -
II.	Effect of the pad layout: metal density, presence of vias	131 -
II.A.	Failure observed after wire bonding tests	132 -
II.A.1.	Pad description	132 -
II.A.2.	Pull tests	133 -
II.A.3.	Shear tests.....	136 -
II.A.4.	Discussion	137 -
II.B.	Nanoindentation tests on CMOS 120nm.....	139 -
II.B.1.	Nanoindentation with Berkovich tip	139 -
II.B.2.	Nanoindentation with cube corner tip	139 -
II.B.3.	Discussion: wire bonding versus nanoindentation	144 -
II.C.	Nanoindentation on CMOS 45 nm pads	145 -
II.C.1.	Pad description	145 -
II.C.2.	Measurements with Berkovich tip.....	146 -
II.D.	Discussion on the pad layouts effect	148 -
III.	Metal copper layer design	148 -
III.A.	Pad layouts description.....	148 -
III.B.	Pad integrity tests	149 -
III.C.	Nanoindentation	151 -
III.C.1.	Berkovich tip	151 -
III.C.2.	Cube corner tip	151 -
III.D.	Numerical results.....	152 -
III.E.	Discussion	153 -
IV.	Nanoindentation and modelling to reproduce EWS and wire bonding.....	155 -
V.	Conclusions	156 -
References	158 -

In the previous chapter, the tests and process parameters during the EWS (overdrive and number of passes) and the wire bonding (contact force and bonding power) have been shown to induce failures in the pad structure. Actual related tests and modelling have been achieved to obtain these results. The nanoindentation and some methods of observation and analysis have been presented in the chapter two. During an indentation test, a compressive load is applied to the pad, which can allow a cheap and fast reproduction of the EWS and assembly loading conditions.

In this fourth chapter, a top-down approach of the various scales set up in the pad volume is performed. Indeed, several parameters can be modified to improve the mechanical robustness of the pads. Hence, the thickness of the aluminium and pad open oxide layers is initially investigated. Afterwards, the effects of the metal density and the presence of vias and pad open vias are questioned. In a third section, several designs for the top copper level (lines and slots) are considered. For each of these scales, several structures are evaluated and general trends are given, but also a more general methodology to analyze the pads through experimental and numerical techniques.

For both probing and wire bonding, actual tests and characterizations are performed to provide baselines for further comparison. These trends are deduced according to the crack and peeling rates. Therefore, for each layout, a comparison is performed between the results of the actual tests, these of the nanoindentation tests and the EWS simulations. For both techniques, some adequate indicators are described to get relevant rankings according to the mechanical robustness of the pad structures. The objective is to establish correlations between the die process and testing, the experimental and numerical techniques. Finally, a full methodology for the technology development and the crisis resolution at the pad level is obtained.

I. Stacking influence: thickness of the top layers

The thickness of the layers is often defined for a technology node. However, on some particular products, while failure issues are faced, this parameter can be modified. Hence, in this section the effect of the thicknesses of the aluminium capping layer and the pad open oxide one just below is investigated. The aluminium thicknesses evaluated are $0.88\mu\text{m}$ and $1.2\mu\text{m}$, and the oxide thicknesses are $0.5\mu\text{m}$ and $0.7\mu\text{m}$. For clarity purposes, the lowest value is often noted L and highest H.

The analyses are performed on CMOS 120nm pads, following all the tests conditions presented in the section I.B. of the third chapter. From bottom to top, this structure is composed of four copper layers (M1 to M4), with Inter-Metal Dielectric (IMD) layers in-between, and then one pad open oxide and finally the aluminium capping layer, as depicted by the Figure IV.1.

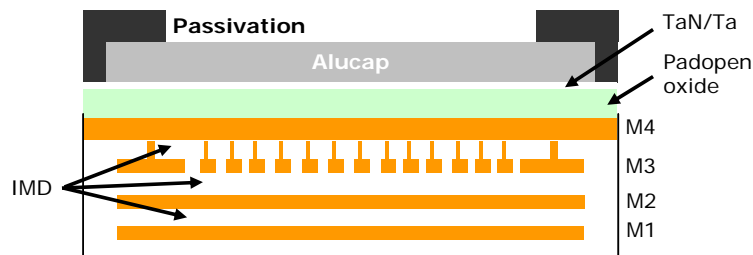


Figure IV.1. Schematic cross section of the evaluated CMOS 120nm pads.

I.A. Failure analyses after EWS

The cracks are firstly observed after pad integrity tests, with the help of optical microscopy, and then using more dedicated techniques to gain insights on the location and root causes of the cracks.

I.A.1. Crack occurrences after pad integrity tests

This study follows the one started in the section I.B.2 of the previous chapter. The failures are located in the pad open oxide layer, as viewed in the Figure IV.2. Multiple cracks are observed in the oxide layer, and the copper is eroded due to the use of chemicals that infiltrates through the fractures (see section II.B. on the chapter 1 for additional details). All the failures are located under the probing area.

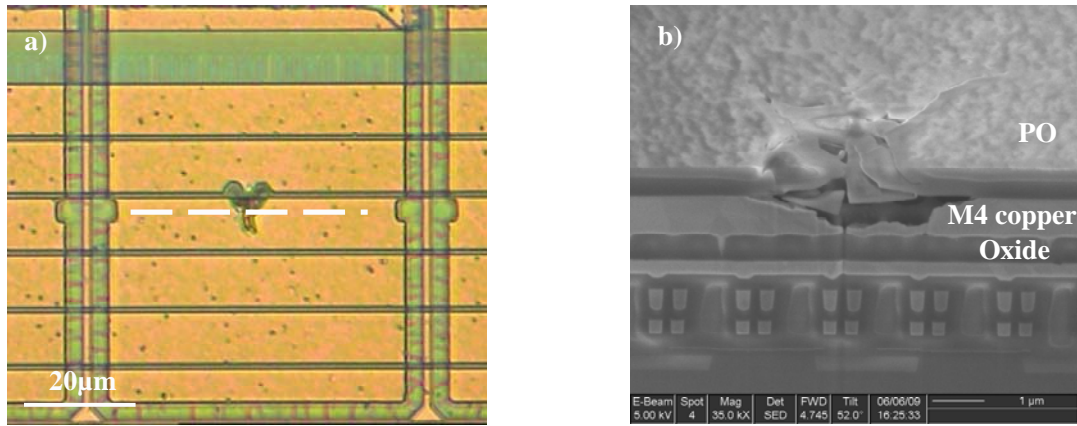


Figure IV.2. Optical microscopy top view of a pad after pad integrity test (a) and SEM cross section of the corresponding pad after FIB milling along the white dashed line (b).

From these observations, the crack occurrence for various tests and stacking parameters is estimated, as depicted in the Figure IV.3. The crack occurrence is found null for the stronger stacks, and can reach up to 80% of the pads for the weakest one. It is demonstrated that the increase of the thicknesses, whether aluminium, oxide or both, induces a reduction of the failures observed in the pad. It remains valid whatever the test conditions.

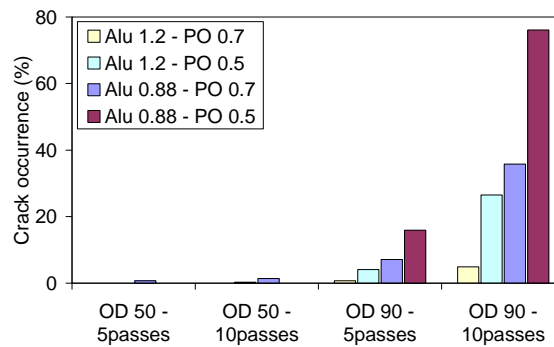


Figure IV.3. Crack occurrence for various EWS tests and pad design conditions.

Moreover, a ranking according to the mechanical robustness of the structures, after EWS testing can be deduced:

$$\begin{array}{|c|} \hline \text{Alu} - \text{L} \\ \hline \text{PO} - \text{L} \\ \hline \end{array} < \begin{array}{|c|} \hline \text{Alu} - \text{L} \\ \hline \text{PO} - \text{H} \\ \hline \end{array} < \begin{array}{|c|} \hline \text{Alu} - \text{H} \\ \hline \text{PO} - \text{L} \\ \hline \end{array} < \begin{array}{|c|} \hline \text{Alu} - \text{H} \\ \hline \text{PO} - \text{H} \\ \hline \end{array} \quad (\text{IV. 1})$$

Optical microscopy is very efficient to count the cracks for a large amount of pads. Therefore, the use of chemicals to reveal the cracks does not allow to observe very accurately the crack path in the oxide layer and to further understand some interactions with the design of the layers underneath.

I.A.2. 3D reconstruction of the cracks under the pads

In order to get a deeper knowledge about the failures, the volume of the pad is reconstructed in 3 dimensions by means of the FIB/SEM observation whose method has been described in the second chapter.

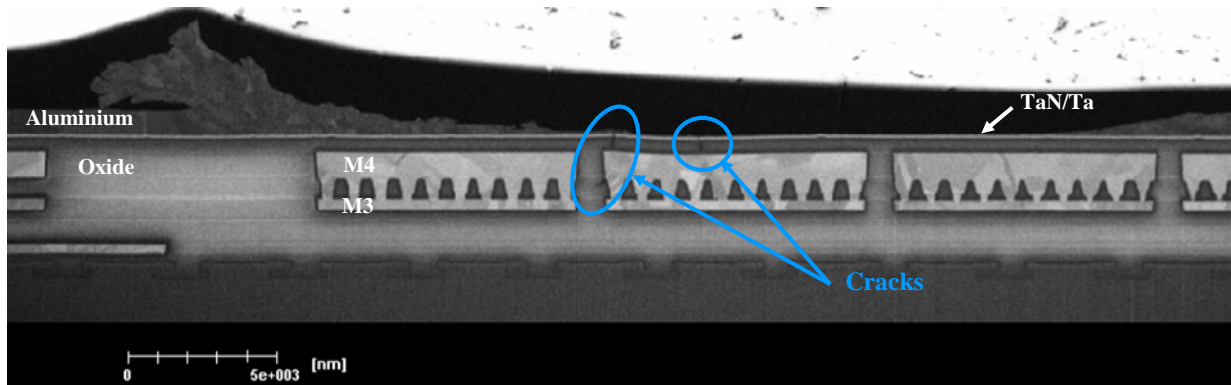
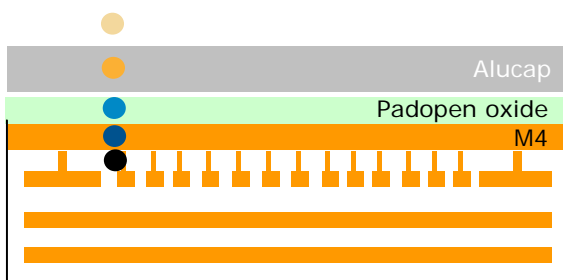
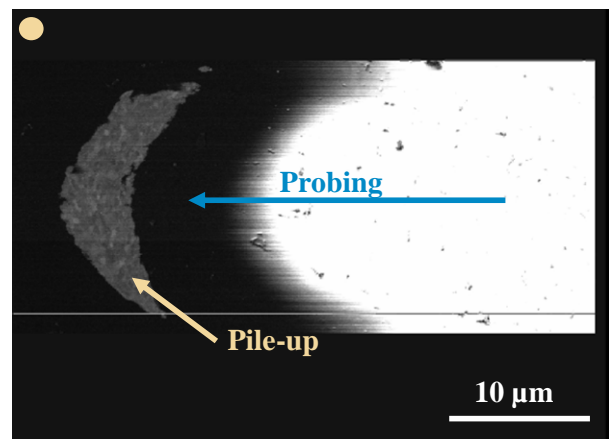


Figure IV.4. SEM cross section of the pad (true image).

The Figure IV.4 shows a cross section of the centre of the probe mark, in parallel of the probe path. Similarly to previous observations, high pile up is observed in the front of the probe mark (see the left of the aluminium). Some major cracks are encountered in the pad open oxide alike the observations of the pad integrity tests. These cracks are located under an area where no or little aluminium remains on the pad. Moreover, smaller cracks, located at geometrical singularities are also observed in oxide layers deeper into the pad.



Schematic side view of the pad



(a) Above the pad

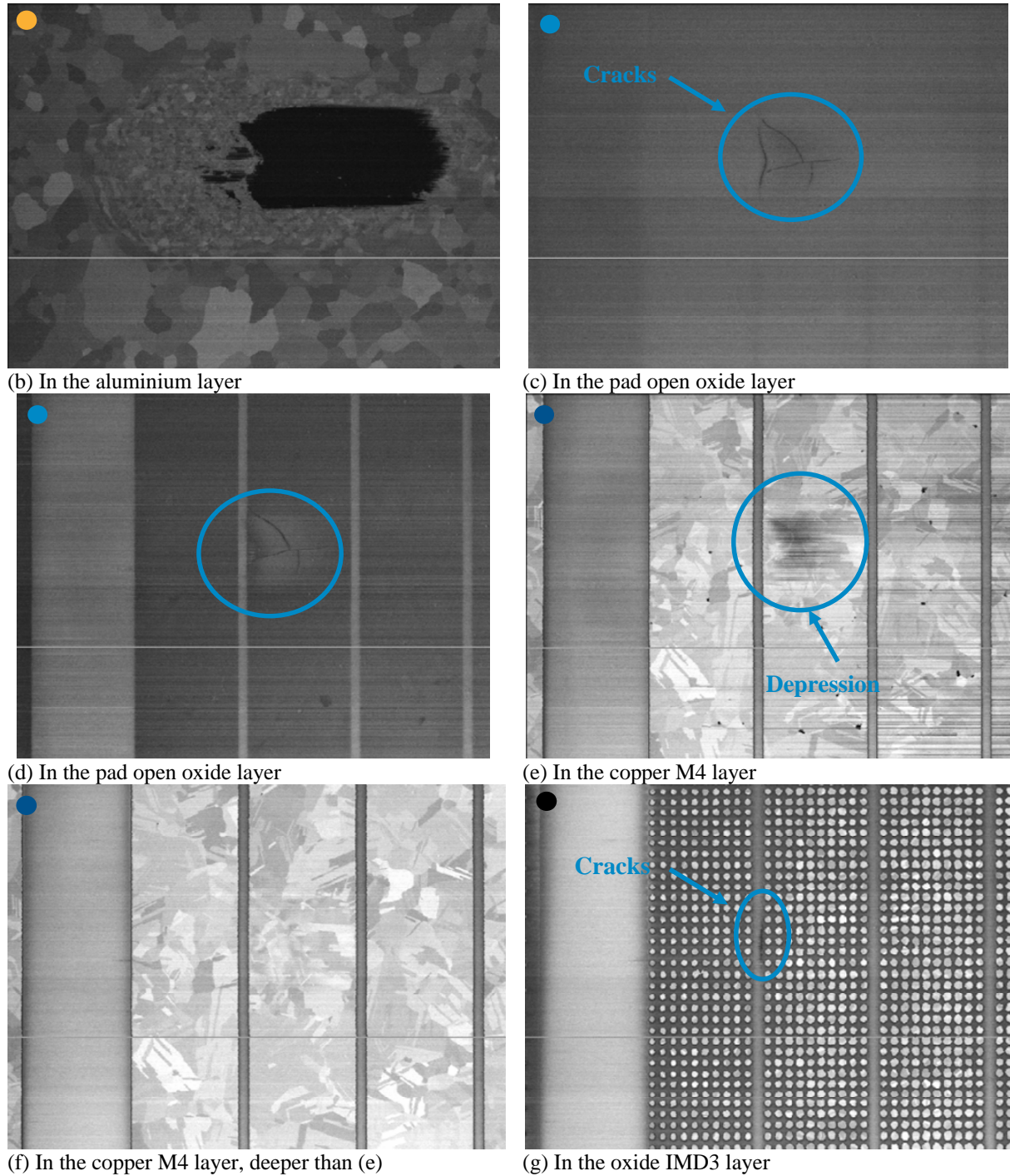


Figure IV.5. Top-down views of the reconstructed pad.

The 3D reconstruction of the pad allows observing and featuring the pad in each direction, and thus the oxide layers containing the cracks. The Figure IV.5 shows several reconstructed top views of the pad from above its surface down to the IMD3 layer. The view (b) shows the finer grains observed in the probing area due to the work hardening of the aluminium during the test. A wide crack (roughly $6.3\mu\text{m} \times 5.4\mu\text{m}$) in multiple directions is seen in the pad open oxide layer (c). Additionally, some close interactions between the crack path and the pad design are shown (d), since fracture path is found to be affected by the copper layer below the oxide. A depression is observed in the top copper layer (e) and has to be related to the plastic deformation of the copper during the probing. Then, still deeper into the architecture, an almost vertical crack is observed in the oxide of the IMD3 layer, from the geometrical

singularities of the M4 copper layer (g). The copper from the third level does not show significant plastic deformation and seems unaffected by the EWS test.

These cracks and deformation in the structure underline the high stresses undergone by the interconnections during EWS tests and provides guidelines for further pad designs. Hence, the aluminium, a ductile material, acts as a protective film for the brittle oxide layers of the pad. Aluminium is a stress buffer and the control of its remaining thickness after the test can be a good indication of the failures probability. The top copper layer is plastically deformed during the test and has an influence on the crack path. The design, and particularly the edges, and consequently the geometrical singularities, of the layer under the PO appear to be the site of either initiation or end of propagation of the cracks. Yet, a closer comparison between the views (c), (d) and (e) of the Figure IV.5 does not show any correlation between the crack path and the M4 copper grain size or distribution.

Additionally, this reconstruction provides confirmations of the assumption and results already demonstrated. Thus, the crack in the PO layer is located close to the half length of the probe mark. As a consequence, in that particular case, the impact of the probe on the pad cannot be suspected to be the cause of the failures. On the numerical model, the TaN/Ta layer has not been considered specifically. The relevance of the simplification is confirmed by the cracks and deformation of this layer, which are similar to the PO oxide, as depicted in the Figure IV.4.

1.B. Resistance to compression: nanoindentation tests

The nanoindentation corresponds to a similar loading condition as EWS test. Indeed, only compressive loading is applied to the pad. Since the experiments to calibrate the numerical model have shown that the vertical forces are higher than the lateral ones, it does not appear to be limiting. Hence, two results are expected while using the nanoindentation to mimic probing. Firstly, the ability to reproduce the cracks observed in the pad open oxide layer, and secondly, the faculty to compare and rank the structures according to their mechanical robustness and obtain the same results as those described in the previous section. The nanoindentation tests with Berkovich and cube corner tips are described for a given pad layout (copper M4: lines) to keep this parameter constant.

As described in the chapter 2, the nanoindentation raises several issues for the measurement of mechanical properties on multi-layers stacks. In our case, the aluminium, which is prone to very important pile-up, and the objective to obtain cracks in the oxide thin layer, harden the pad study. At the same time it reduces the reliability of the hardness and the Young's modulus measurements due to the multilayer system. As a consequence, both Berkovich and cube corner tips are used during the testing for their distinct capabilities.

1.B.1. Cube corner tip

1.B.1.a) Load determination and indent observations

In order to evaluate the load necessary to penetrate deep enough into the surface, i.e. in the aluminium and oxide layers, some tests with increasing loading are performed, as shown in Figure IV.6. The nanoindentation in the oxide layer is required to generate crack within, and to try to reproduce the failures observed during probing. As expected, all the curves overlap each other in the loading part. A loading of 15mN is selected to penetrate through the oxide film. Then, for reproducibility purposes, 35 indents are performed for each structure.

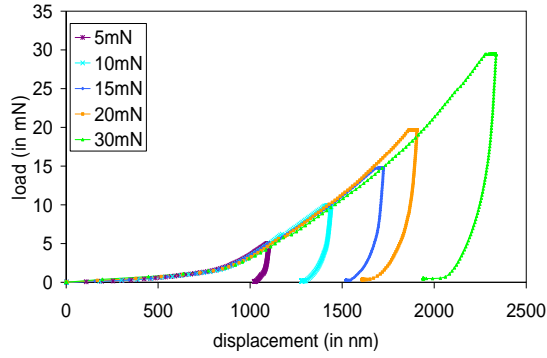


Figure IV.6. Load-displacement curves for various loading conditions.

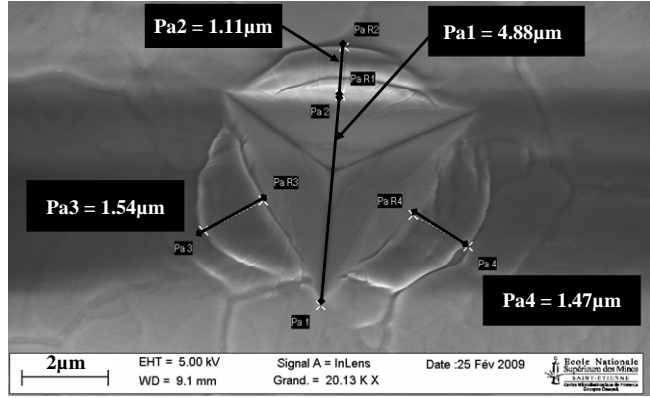


Figure IV.7. SEM top views of the indent made by a cube corner tip on the alucap.

The cube corner tip is very sharp. As a consequence, the size of the indent left after the indentation is very low. For an indent reaching a depth of $2\mu\text{m}$, the triangle height of the projected area remaining is below $5\mu\text{m}$. The pile up is very important for such tests on aluminium. More precisely, for this loading condition, the width of the pile up is found between 1.1 and $1.6\mu\text{m}$, as shown in Figure IV.7. Its height is measured during the FIB milling and SEM observations of samples, and estimated to be in the order of $0.6\mu\text{m}$ for the sample with $1.2\mu\text{m}$ thick aluminium, which is consequent.

I.B.1.b) Analysis of the load-displacement curves

The load-displacement curves for the stack with lower values of aluminium and PO thicknesses are depicted in Figure IV.8. A quite good reproducibility of the nanoindentation with cube corner tip is observed for both the loading and unloading stages. Firstly, during the loading step, a clear slope break is observed between the aluminium and the oxide layers. Indeed, the tip goes through the aluminium very easily since only about 2mN are sufficient, and in the oxide part, the penetration of the tip is much harder. These two films have close Young's moduli, which are respectively 69GPa and 63GPa for the aluminium and the SiO_2 . However, the hardness of the aluminium layer is close to 0.6GPa , while this of the oxide is close to 7.7GPa .

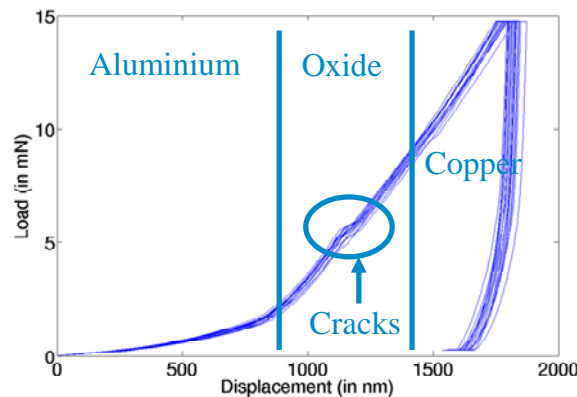


Figure IV.8. Load-displacement curves for the experiments on the structure with lower values of aluminium and oxide.

Moreover, a small discontinuity is observed in all the curves in the oxide part. This discontinuity is the signature of cracks in the structure. This failure is vertical and located in the oxide layer, as shown in the Figure IV.9. However, the cracks observed in the pad open oxide layer during the indentation with the cube corner tip can not be easily and directly

compared to those due to the EWS tests, mainly because of the sharpness of the tip that highly damages the oxide layer. The only similarity is its verticality through the whole thickness of the film, which also indicates good adhesion energy between the oxide and the surrounding TaN/Ta and copper layers.

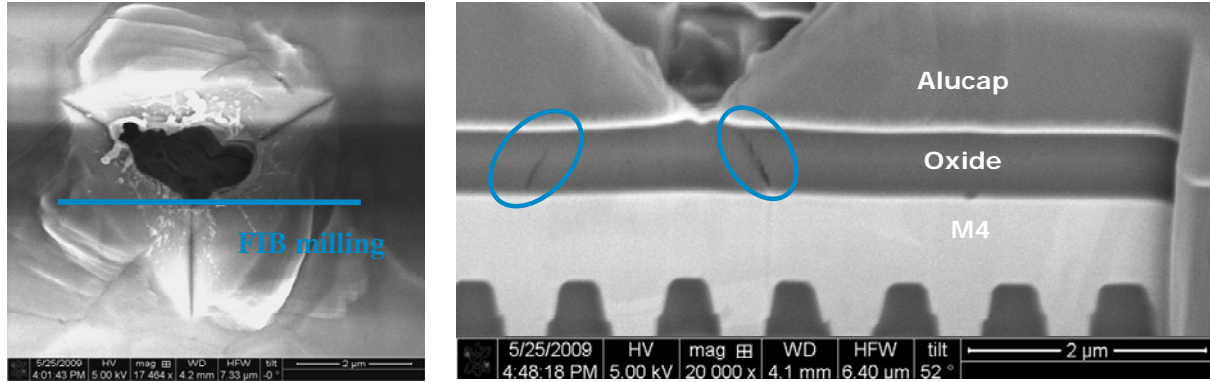


Figure IV.9. SEM tilted cross section of the pad after FIB milling.

At the beginning of the unloading curve, an almost vertical straight line is observed, and then a low elastic recovery occurs. In the literature, several studies [Doerner, 1986; Oliver, 1992] on bulk aluminium have shown that aluminium depicts very low elastic recovery during the unloading stage, i.e. an almost vertical unloading curve (h_f/h_{max} close to 1) is obtained. The case of the copper is more widely investigated in further sections, but it also shows an important plastic deformation and a low elastic recovery. However, it seems hard to attribute this behaviour to the oxide film only. A complex mechanism due to the stacking of the three layers is probably involved, and release of stressed layers might have an effect as well. Then, the load-displacement curves are averaged for the various stacks, and the results are depicted in the Figure IV.10.

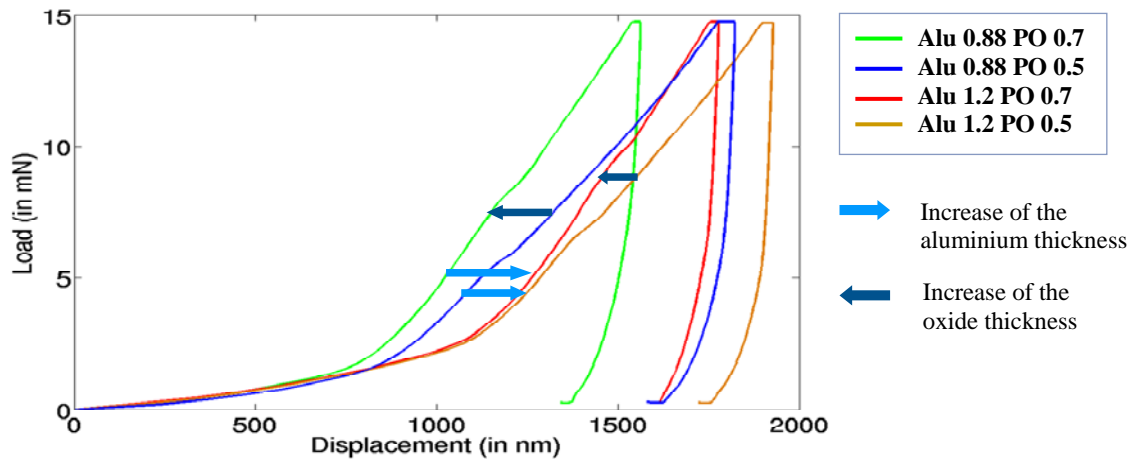


Figure IV.10. Average load-displacement curves for the various stacks.

The increase of the thickness of the aluminium (blue arrows) is shown to increase the depth reached by the tip at the end of the loading. In the literature, Beegan et al [Beegan, 2005] found a similar effect while increasing the thickness of the top layer, even though the test have been performed with copper on oxidised silicon substrates instead of aluminium on top of the pad.

On the contrary, the decrease of the oxide thickness increases the maximum displacement into the surface. The depth at the end of the loading can be related to the ability of a structure

to resist to vertical compression. Hence, a ranking can be established according to this parameter:

$$\begin{array}{|c|} \hline \text{Alu} - H \\ \hline \text{PO} - L \\ \hline \end{array} < \begin{array}{|c|} \hline \text{Alu} - L \\ \hline \text{PO} - L \\ \hline \end{array} < \begin{array}{|c|} \hline \text{Alu} - H \\ \hline \text{PO} - H \\ \hline \end{array} < \begin{array}{|c|} \hline \text{Alu} - L \\ \hline \text{PO} - H \\ \hline \end{array} \quad (\text{IV. 2})$$

This ranking does not match with the results from the actual EWS tests. Indeed, the effect of the aluminium thickness is not well taken into account since for the highest thicknesses a shift is already observed at the end of the aluminium layer due to its low capability to resist to compression with a sharp material. Thus, in that particular case, only the effect of the oxide layer is well reproduced. This is characterized by the curves overlapping in the aluminium domain.

I.B.1.c) Determination of a novel indicator to characterize the cracks

Hence, it appears necessary to find relevant indicators that characterize all the thicknesses effects. Until now, the very small cracks have been only mentioned. A comprehensive study of their corresponding depth and loading of appearance could provide additional characterizations on the robustness of the structures. Indeed, their observations on the load-displacement curves are quite difficult because most of them are very small, and usually, they are related to discontinuities lower than 50 nm. In order to emphasize them, $h \cdot dh/dp$ is plotted versus the displacement or the load, as shown in the Figure IV.11. The inverse of the slope dh/dp correspond to the ability to resist to the tip penetration. The multiplication by the displacement into the surface underlines the effect of the small plateau related to the crack. From these curves, high values correspond to an easy sink in the structure, and as a consequence to the aluminium layer and to the cracks. The value of $h \cdot dh/dp$ is almost constant in the oxide layer.

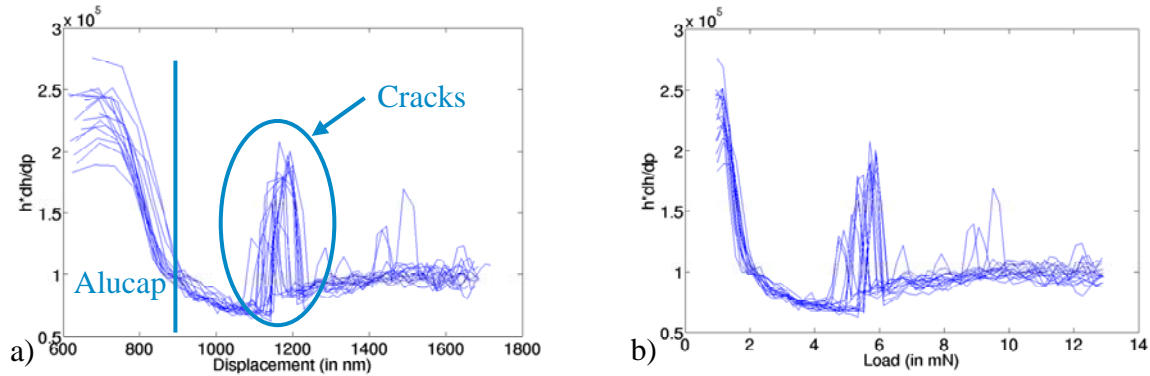


Figure IV.11. $h \cdot dh/dp$ curves as a function of the displacement (a) and the load (b) for the stack with lower aluminium and oxide thicknesses.

Thus, from these curves, the maximum value of $h \cdot dh/dp$ for each pad structure is used to extract the critical loads and depths related to the cracks, which are summarized in the Figure IV.12. The depth of the crack seems mainly dependant on the thickness of the aluminium layer, and in both cases, is found to be close to the middle of the half of the oxide thickness. For the crack load, even if quite high uncertainty is obtained for the values, a clear trend between the layers is obtained. Increasing the thicknesses of the top two layers induces a delay in the appearance of the cracks load because higher crack loads are necessary.

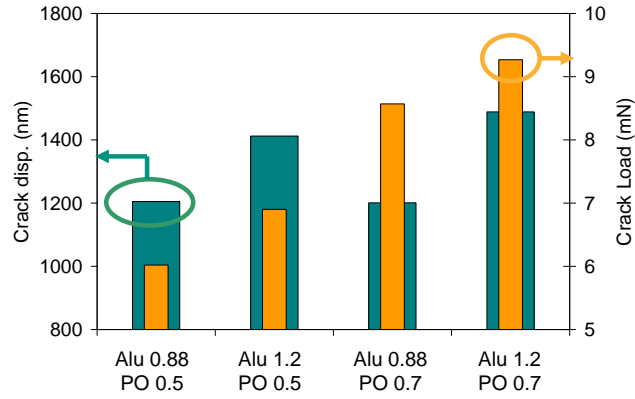


Figure IV.12. Cracks load and depth for the various pad stacks.

The variation of the aluminium thickness does not invalidate the trends and the various stacks can still be compared. Thus, considering that the pad design below the pad open oxide is exactly the same for all the structures, the differences of slopes in the oxide part of the curve should only be related to effects of these two layers. As a consequence, the increase by about 1kN/m in each case while increasing the aluminium thickness can be directly related to the ductile properties of the aluminium, that act as a buffer and to an additional force that is required to overcome this thickness increase. The increase of the oxide thickness induces an increase (about 4kN/m), which depict the increase of its resistance to compression. All the results aforementioned are summarized in the Table IV.1.

	Maximum displacement (nm)	Crack depth (nm)	Crack load (mN)	Slope of the curve in the oxide (kN/m)
Alucap 0.88μm PO 0.7μm	1542 (57)	1201 (57)	8.6 (1.3)	18.2
Alucap 0.88μm PO 0.5μm	1778 (20)	1205 (105)	6.0 (1.3)	14.3
Alucap 1.2μm PO 0.7μm	1758 (46)	1489 (59)	9.3 (1.1)	19.2
Alucap 1.2μm PO 0.5μm	1904 (43)	1412 (43)	6.9 (0.5)	15.2

Table IV.1. Extracted maximum displacement, crack depth and load and slope in the oxide layer and related standard deviations from the nanoindentation curves for a given pad layout (for M4: lines).

Finally, considering all the results described, a ranking according to the resistance to compression and fracture can be established between the various stacks:

$$\begin{array}{c}
 \text{Alu} - L \\
 \hline
 \text{PO} - L
 \end{array}
 <
 \begin{array}{c}
 \text{Alu} - H \\
 \hline
 \text{PO} - L
 \end{array}
 <
 \begin{array}{c}
 \text{Alu} - L \\
 \hline
 \text{PO} - H
 \end{array}
 <
 \begin{array}{c}
 \text{Alu} - H \\
 \hline
 \text{PO} - H
 \end{array}
 \quad (IV. 3)$$

1.B.1.d) Evaluation of E and H

For the indentation at 30mN, 15 indents with the CSM mode have been performed on the stack with low aluminium and pad open oxide thicknesses. Using the Oliver and Pharr method, the hardness and Young's modulus of the structures versus the displacement into the structure are calculated and shown in Figure IV.13 and Figure IV.14. Substrate effects for the values in the aluminium layer are clearly observed on both curves. Considering the 10% limit of the film thickness to avoid substrate effects, the hardness and an elastic modulus are respectively 0.38GPa and 70.7GPa. The former is quite low, usually, the hardness of the

aluminium layer is found between 0.5GPa and 0.7GPa, depending on the process used for the layer deposition. The hardness in the alucap film increases in the layer (i.e. until the 880nm depth) due to the harder oxide below. The substrate effect is clearly observed on both curves. For the Young's modulus, the presence of the copper, whose modulus is about 130GPa, increases the values expected for the aluminium and pad open layers.

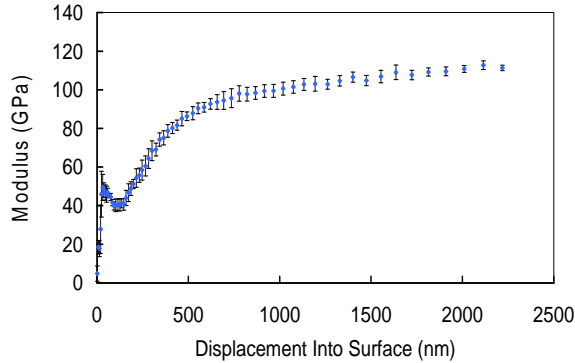


Figure IV.13. Young's modulus as a function of the nanoindentation depth.

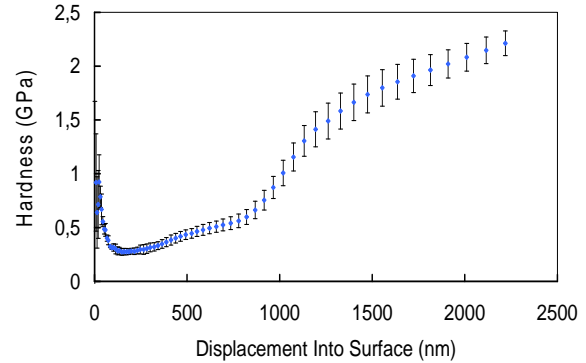


Figure IV.14. Hardness as a function of the nanoindentation depth.

Hence, even if the hardness in the oxide layers can also be supposed to be a good indicator to weight the mechanical robustness of the structures, the previous indicators (maximum displacement, crack load and depth) fulfil the requirements and avoid the mistakes related to the Oliver and Pharr method on multi-layers with high pile up of the top one. Additionally, for the calculation of the material properties of distinct aluminium thicknesses, the pile-up and the projected area are probably different, and so do the deduced hardness and Young's moduli. Indeed, the film-substrate interactions are different since the strains are confined in the aluminium layer.

I.B.2. Berkovich tip

I.B.2.a) Measurements and results

In a second part, some tests are made with the Berkovich tip, which is more commonly used. For each structure, 15 indents are performed for reproducibility purposes. The load applied is set to 200mN. This is about 10 times higher than the load applied during the cube corner experiments, which is expected due to the volume-depth relation of both tips (see section II.A.2. of the second chapter). The indent left by a Berkovich tip in the pad is shown in Figure IV.15 (a) and (b). The mark shows the 3 sides of the pyramid, and for such loading conditions, the triangle height of the projected area is about 16.4 μ m. It is about 3 to 4 times larger than the indent left by a cube corner tip for a similar indenting depth. The aluminium is also clearly plastically deformed during the test, and some pile-up is observed on the edges. The maximum width of the heap is between 2.4 μ m and 3 μ m in that particular case.

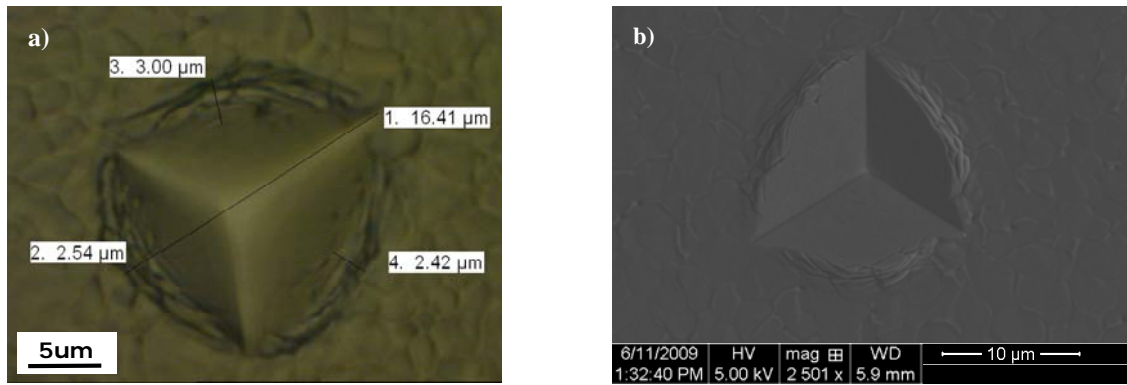


Figure IV.15. Optical microscopy (a) and SEM (b) top views of an indent on the pad made with a Berkovich tip.

The load-displacement curve resulting from such tests is depicted in Figure IV.16. A good reproducibility is obtained for the set of measurements. The loading curve follows a power law equation, and no clear slope variation is observed at the interfaces between the layers. The tip penetrates through the aluminium, then through the pad open oxide, and at the end of the loading reaches the copper layer. No discontinuity in the oxide part can be clearly observed on the load-displacement curves. During the unloading stage, initially, the curve is linear, and then similarly to the cube corner tip, a small elastic recovery is observed.

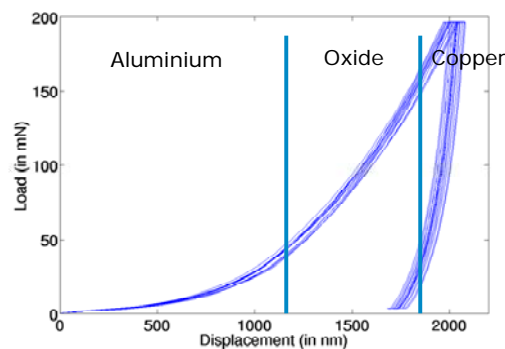


Figure IV.16. Load-displacement curve for nanoindentation using a Berkovich tip (aluminium 1.2μm, PO 0.7μm).

The depths reached at the end of the loading are shown in the Figure IV.17. Increasing the thickness of the aluminium induces an increase of the displacement into the surface. As mentioned previously, this can be explained by the material properties of the aluminium compared to these of the oxide. The difference in aluminium thicknesses between the stacks is 320nm. After the nanoindentation with the Berkovich tip, the shift is only about 180nm (see Figure IV.17). Hence, it underlines the protective effect of the aluminium.

No clear trend can be established for the role of the thickness of the oxide layer, since the standard deviation is higher than the differences between the structures. Moreover, since no discontinuities are observed in the load-displacement curves, it seems that no other information can be obtained.

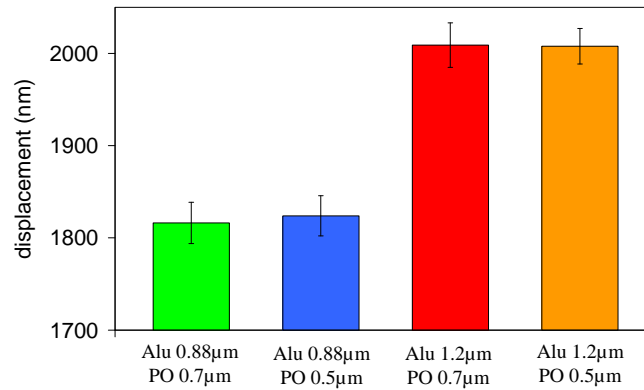
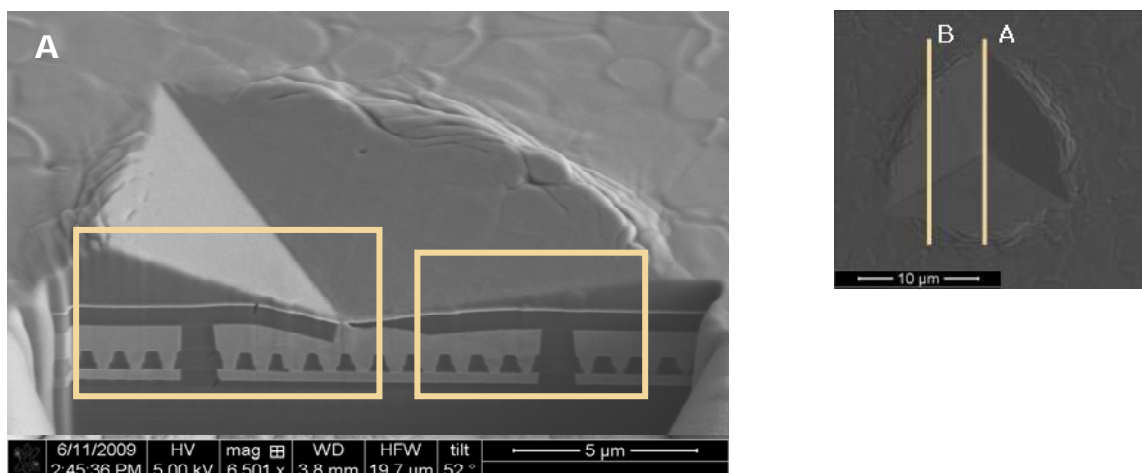


Figure IV.17. Maximum displacement into surface for the various stacks (pad layout with lines).

I.B.2.b) SEM analysis of the sample after FIB milling

On a Berkovich indented pad, some FIB millings followed by SEM views are performed. The observation along one edge of the indenter for the thickest stack (aluminium and pad open thicknesses are respectively 1.2µm and 0.7µm) is shown in Figure IV.18. The pile up reaches about 0.9µm. After the indentation, some aluminium, even in the middle of the indent is observed on the tantalum layer, even if very thin. In the structure some cracks can be observed in the PO oxide and in the IMD layers. The former fracture is located under the indent, in a location where oxide is highly bent. The failure propagates in the TaN/Ta layer as well, which underlines once more that this layer have quite similar behaviour than the oxide below. Contrary to the previous cases, the failure does not go through the whole oxide thickness. In that particular case, the oxide from the pad open layer bends and breaks, since along the soft edge of the indenter (SEM view on the right), no crack is observed. Additionally, a correlation may exist between the crack's location and the indenter geometry since it appears where the oxide starts to bend. It can also explain the appearance of the crack in the layer during cube corner nanoindentation tests, because, the oxide has to be bent enough prior to fail. The initiation of the crack from the top of the oxide layer is related to the most stressed and bent area. Then, the failure does not propagate through the whole thickness of the layer due to the plastic deformation of the M4 copper. The latter cracks are located in IMD2 and IMD3 at the geometrical singularities.

The copper is plastically deformed during the indentation. Then during the unloading stage the release of the load is related to an extrusion of the compressed copper, explaining probably the shape of the unloading curve.



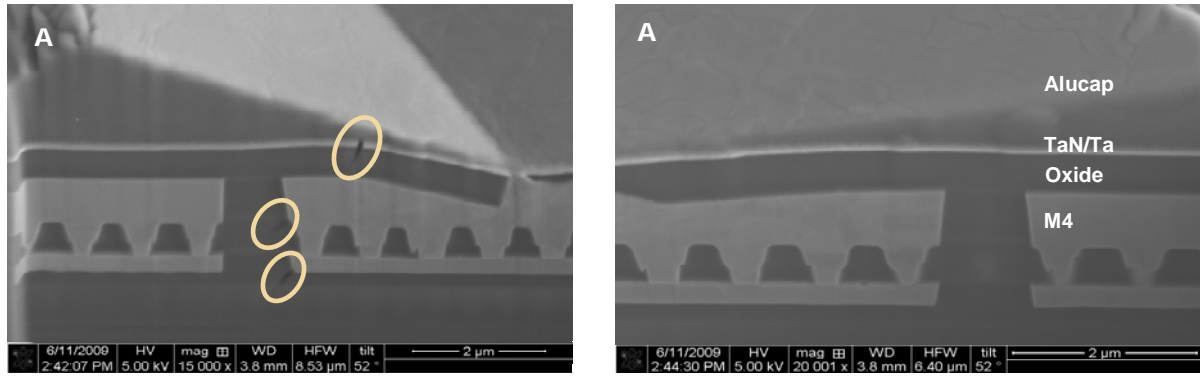


Figure IV.18. SEM tilted cross sections after FIB milling in location A of the pad with highest thicknesses of aluminium and oxide pad open.

The failures in the oxide layers propagate along the indenter's side, as depicted in the Figure IV.19. The behaviour of the copper, more particularly its plastic deformation during the loading stage and extrusion during the unloading one are emphasized as well. The extrusion seems to take place along the edges of the indent. In the literature, Lee and co-workers [Lee, 2007] also observed some cracks in the oxide layer with copper extrusion within the pad during nanoindentation on pads.

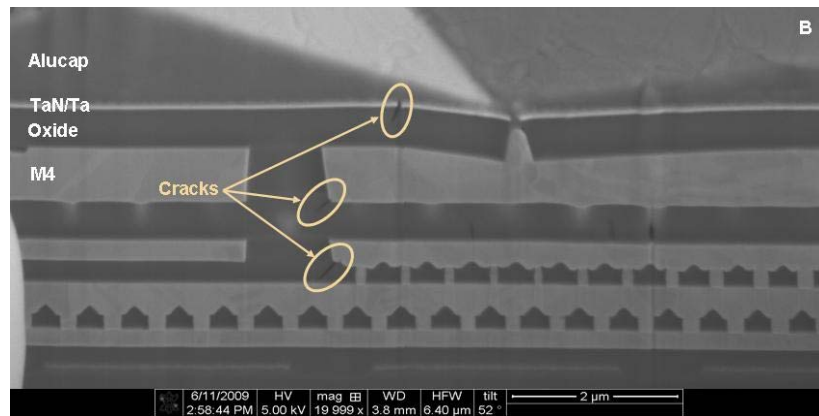


Figure IV.19. SEM tilted cross sections after FIB milling of the pad in the location B.

On the load-displacement curves shown in the previous section, no discontinuity has been clearly observed on the curves. A closer look to the curve corresponding to the FIB/SEM observation, using the h^*dh/dp indicator is shown in Figure IV.20 (a) and (b). On that plot, a very small peak is seen in the oxide region at 1330μm. This small peak underlines that these cracks are related to very small released energy.

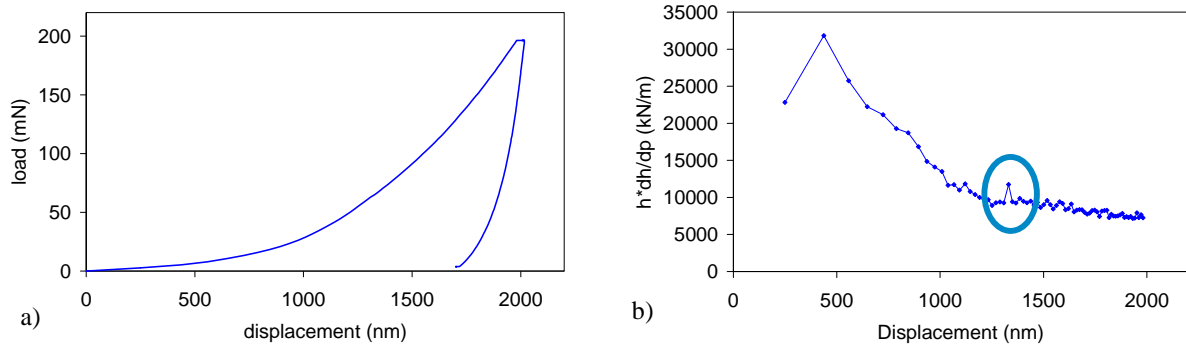


Figure IV.20. FIB/SEM corresponding load-displacement curve (a) and h^*dh/dp versus displacement plot (b).

These failures in the pad structure have to be closed to those observed in the pad during the EWS tests, depicted in the Figure IV.4. Hence, the nanoindentation with the Berkovich tip is very interesting in its ability to reproduce the cracks observed during actual probing.

Finally, nanoindentation with cube corner and Berkovich tips shows some features of the EWS, and has proven its efficiency to evaluate pads, even if both indenters have advantages and drawbacks.

I.C. EWS numerical model: stress analysis

For this study, the OD is fixed to 100 μ m, which corresponds to the harshest probing condition, and as a consequence to the highest stresses undergone by the pad. The description of the pad has been made in the section III.A.1 of the chapter 2. In the section I.B.3 of the chapter 3, it has been demonstrated that the compressive and Von Mises stresses are maximum at the end of the loading stage, during the plateau used to perform the electrical measurements. As a consequence, the study focuses on these steps. The pad and the probe tip during one corresponding substep are shown in the Figure IV.21. For clarity, all the materials are translucent and the iso-stress curves are plotted in the pad only. The maximum is found in the aluminium layer, just under the contact location.

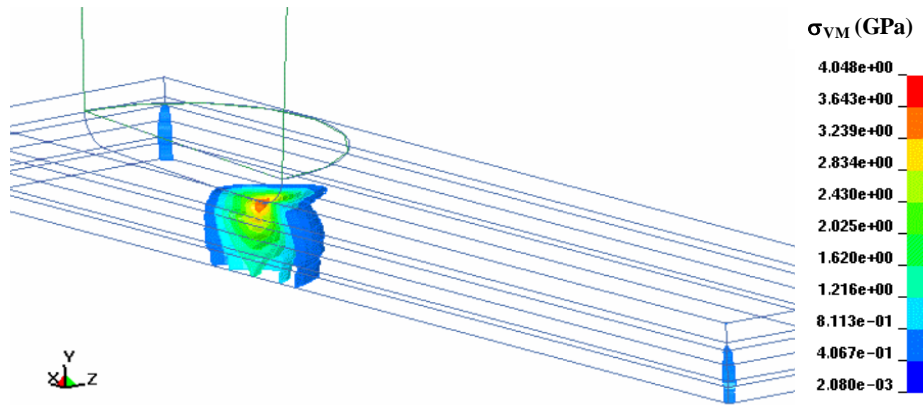


Figure IV.21. Simulated Von Mises iso-stress in the pad.

Then, the compressive and Von Mises stresses are collected on top of each layer, which correspond to the maximum reached stress. As confirmed by the Figure IV.22, the maximum stress is located in the aluminium, and then decreases while reaching deeper layers under the surface of the pad.

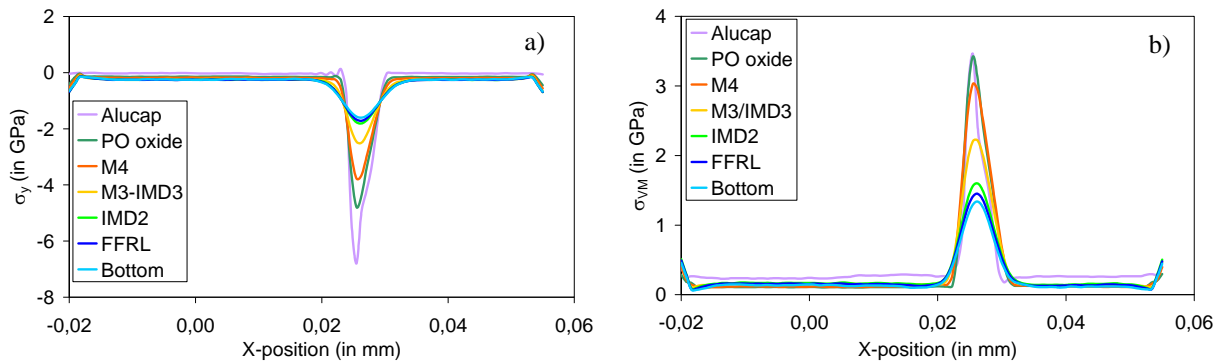


Figure IV.22. Compressive (a) and Von Mises (b) stresses in the various layers along the pad length for the thinnest structure.

The maxima of the top four layers are depicted in Figure IV.23. The most interesting layers are those containing oxide since cracks are found within. Hence, the comparison between the structures is made according to the stress values in the PO and in the M3/IMD3 layers. Increasing both the thicknesses of aluminium and pad open oxide is found to reduce the stress in the investigated layers. Since lower stress is undergone by the layer, lower crack occurrence is expected in the structure. Similar trends are obtained analysing the compressive stress. The cumulative and separated thickness effects are evaluated through modelling.

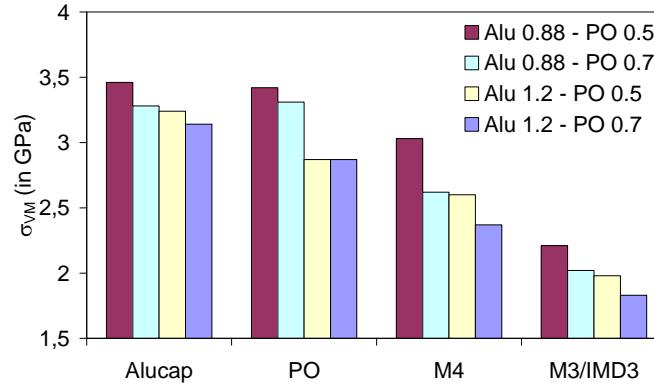


Figure IV.23. Maximum Von Mises stress in the top layers for the various stacks.

Thus, a ranking, according to the maximum Von Mises in the structure, is obtained:

$$\begin{array}{|c|} \hline \text{Alu} - L \\ \hline \text{PO} - L \\ \hline \end{array} < \begin{array}{|c|} \hline \text{Alu} - L \\ \hline \text{PO} - H \\ \hline \end{array} < \begin{array}{|c|} \hline \text{Alu} - H \\ \hline \text{PO} - L \\ \hline \end{array} < \begin{array}{|c|} \hline \text{Alu} - H \\ \hline \text{PO} - H \\ \hline \end{array} \quad (\text{IV. 4})$$

I.D. Discussion and correlation between techniques about the thickness effect

I.D.1. Comparison between techniques to evaluate the thickness effect

In the previous sections, a ranking according to indicators proper to each technique have been achieved. The crack occurrence after the EWS test is the reference (see ranking (IV. 1)). Indeed, increasing the thicknesses of the aluminium or the pad open oxide layers is demonstrated to reduce the crack occurrence in the pad during testing. The ability of nanoindentation and modeling to mimic the effect of the thicknesses of aluminium and PO oxide is investigated.

The stress analysis, based on the Von Mises or the compressive stress made on the numerical model reproduces all the stacking effects and the ranking obtained is the same as the actual test one ((IV. 4). The aggregate effect linked to the variation of both layers is also well depicted.

The nanoindentation with the cube corner tip is able to reproduce the increase of the thickness of the two layers separately. Using only the comparison of displacement into the surface at maximum load is not sufficient, and the analysis of the load and displacements of the cracks induced by the penetration of the tip in the oxide enables us to match the trends (IV. 3). However, the cumulative effect is not obtained due to the differences of material properties between aluminium and oxide (TEOS). The nanoindentation with the Berkovich tip is only able to compare pad with various aluminium thicknesses. Therefore, the same crack locations are found during these indentations as during EWS. Hence, nanoindentation with both tips provides interesting insights for the pad evaluation.

To summarize, for the stacking parameters, the experimental and numerical analyses, and their related indicators set up are efficient to reproduce the results obtained during the EWS test.

I.D.2. Discussion

In this study, increasing the thicknesses of the aluminium and/or the pad open oxide, and a cumulative effect while increasing both at the same time are shown to reduce the failure hazard in the pad. The influence of the aluminium thickness has been expected due to the ductile material properties of aluminium. The aluminium acts as a buffer to reduce the stress undergone by the pad, and is submitted to large plastic deformation during the test. In the literature, Tran et al. [2000] showed that augmenting the aluminium thickness is related to lower issues during EWS. However, due to manufacturing issues, they preferred the mid-size aluminium thickness.

In their papers, Liu and co-workers [Liu, 2005] evaluated the effect of various thicknesses of aluminium and IMD from experiments and modeling. They obtained a reduction of the number of cracks and of the first principal stress while increasing the layer thicknesses, which is in agreement with the results of the current study. In a more recent study on a different pad structure (see Figure IV.24), the same authors [Liu, 2007] evaluated through a stress analysis from their FE model that increasing the thicknesses of the SiN layer lead to a reduction of the first principal stresses in the SiN and dielectric layers. On the contrary reducing the thickness of the copper, the nickel is shown to reduce the stresses in the investigated layers. The thickness of the M2 layer and third dielectric level has no or limited effect on the stresses. Hence, due to its higher Young's modulus and hardness, the copper does not allow a reduction of the stress the same way as aluminium. However, the thick copper and the global stacking of the pad are very different from the studied pad, which explains probably some differences in the results.

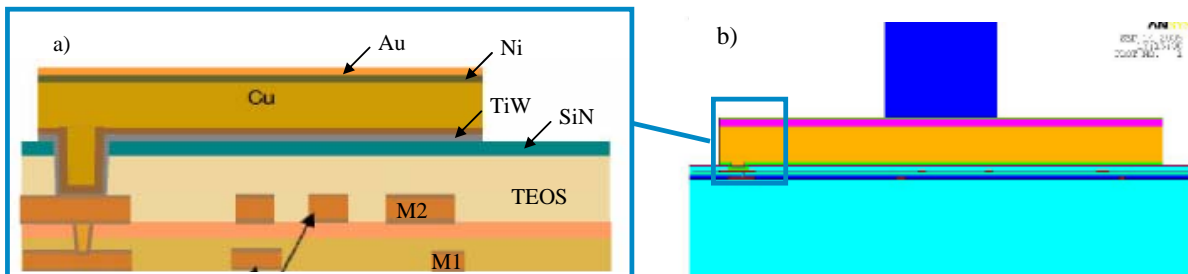


Figure IV.24. Bond pad over active structure (a) and corresponding 2D FE model (b) after [Liu, 2007].

II. Effect of the pad layout: metal density, presence of vias

In the previous sections, the FIB/SEM observations after the probing tests have underlined the protective effect of the copper on the structure. Hence, in this section, the metal density and the presence of vias and pad open vias are investigated. These effects concerns smaller scales than previously described ones. As a matter of fact, the copper density of the layers at the bottom of the stack, just above the silicon is evaluated. It varies from empty to full sheet (100% copper). As a consequence of the design and routing required in the structure, it is sometimes possible to modify the layer's density, depending on the pad technology. It induces a variation of the overall rigidity of the structure, which is expected to have a role on the failure hazard. Moreover, between the copper levels, some vias can be required to ensure the electrical connections. Between the top copper layer, and the aluminium capping, some vias in the pad open layer may be necessary for the same reason.

II.A. Failure observed after wire bonding tests

As part of the large copper bonding DOE aforementioned in the section III of the chapter 3, the effect of the metal density, the presence of vias and vias pad open are investigated. Wire pull and shear tests are made in order to evaluate the resistance to peeling. These structures are quite different to those evaluate in the frame of the EWS mimicking.

II.A.1. Pad description

The technology node for these pads is CMOS 120nm. The copper density of the top (M4) layer is fixed at 80% due to design requirements, hence, only the densities of M1 to M3 are investigated. They are set at either 30% or 80%. The remaining material is some squared TEOS oxide filled holes in the copper plate with a given size for each design. This pad layout is also often referred as slots. The current layout induces different surfaces roughness between architectures with and without pad open vias, as depicted in Figure IV.25 and Figure IV.26. For the structure without PO via, the aluminium is directly on top of the M4 layer (mostly copper). For the one with via pad open, the aluminium is either on M4 copper or on the pad open oxide, which implies the roughness variation. Yet the aluminium thickness is constant throughout the pad. Between the copper layers M1 to M4, the presence of vias is investigated, the pad contains vias between all the layers or none of them.

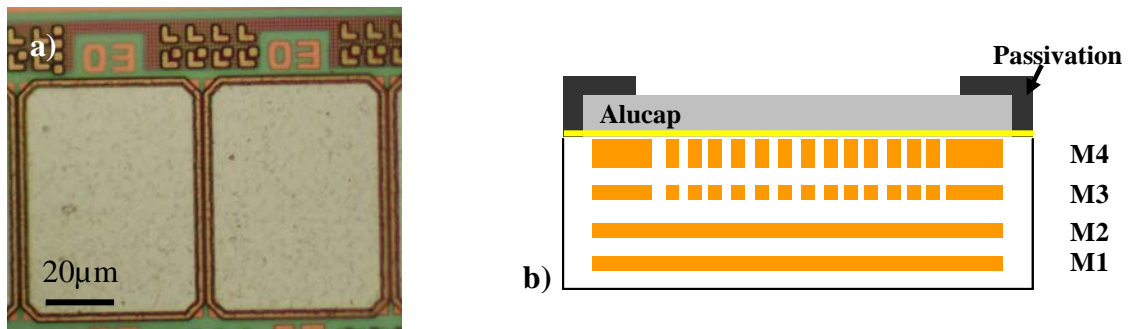


Figure IV.25. Optical microscopy top view (a) and schematic cross section (b) of pads without PO.

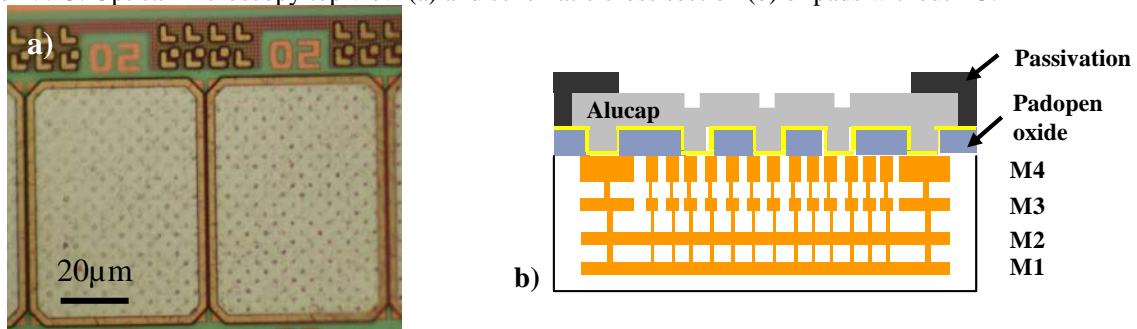


Figure IV.26. Optical microscopy top view (a) and schematic cross section (b) of pads with PO vias.

Pad structure	Metal density (M1 to M3)	Presence of vias	Presence of PO with vias
A	30%	No	Yes
B	80%	Yes	Yes
C	80%	No	No
D	80%	No	Yes
E	30%	Yes	No
F	30%	No	No
G	80%	Yes	No
H	30%	Yes	Yes

Table IV.2. DOE matrix for the study of pad layouts.

The DOE matrix for the study of the design parameters is shown in Table IV.2. Then, for each of these structures, the bonding is achieved and wire pull and shear tests are made.

II.A.2. Pull tests

II.A.2.a) Peeling and neck break occurrences

In order to evaluate with quite good accuracy the rates, 45 wires for each pad layout and each set of bonding parameters are pulled. Hence, 405 wires are pulled for each structure and 3240 wires for the whole DOE.

The main target of the wire pull test is to obtain peeling in order to discriminate the structures according to their mechanical robustness. The effect of the bonding parameters has been already discussed (see section III.B in the chapter 3), thus, only the effect of the design parameters is focussed on. As previously mentioned, both peeling and neck break are obtained. In most of the cases (3142/3240 wires, i.e. 97%), failures at the neck are observed. However, high disparities according to the parameters are seen. Thus, the peeling occurrence is null for most of the pads, but reaches 33% for the pad D with specific bonding conditions (CF = 30mN, BP = 90mW). The number of peeling cases for each condition is shown in the Table IV.3.

CF/BP Pads	30 70	30 80	30 90	40 70	40 80	40 90	50 70	50 80	50 90	Total
A	0	2	10	0	1	6	0	0	0	19
B	0	1	7	0	2	5	0	0	3	18
C	0	0	2	0	0	1	0	0	0	3
D	0	1	15	0	0	12	0	0	1	29
E	0	0	0	0	0	0	0	0	0	0
F	0	0	1	0	0	0	0	0	0	1
G	0	0	0	0	0	1	0	0	0	1
H	0	0	9	0	1	12	0	1	4	27
Total	0	4	44	0	4	37	0	1	8	98

Table IV.3. Number of peeling cases obtained among 45 wires for the various bonding parameters (contact force and bonding power) and layouts.

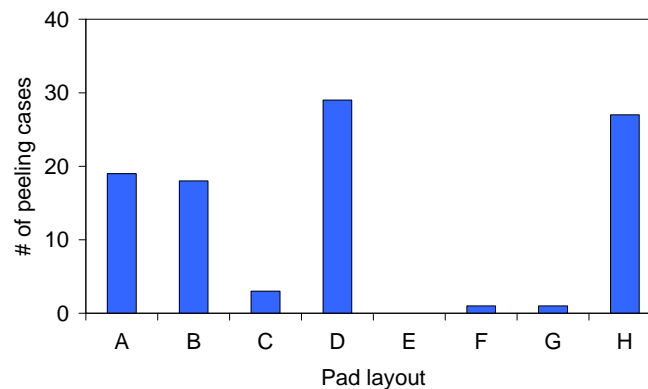


Figure IV.27. Number of peeling cases for all the layouts, independently on the bonding conditions.

The presence of the pad open via is clearly observed to increase the peeling occurrence since the four pads containing this layout show the highest rates, as depicted in the Figure IV.27. The metal density and the presence of via do not modify the robustness of the structure while faced to pull tests. Hence, a ranking according to the mechanical robustness of the pads is obtained:

$$D < H < A < B < C < F \approx G < E \quad (IV.5)$$

II.A.2.b) Observation and analysis of the peeling cases

Then, the observation of the peeling cases is required to locate the failures and try to find a failure mechanism for this kind of pads. Therefore, SEM top views are performed and coupled with EDX analyses to identify the remaining layers.

Similarly to the peeling results, depending on the presence of pad open via, distinct delamination locations are observed. For the pads with pad open vias (A, B, D and H), the peeling is often found between the top copper layer (M4) and the pad open oxide, as shown in Figure IV.28. The EDX analyses show that remaining layers on top of the surface after peeling are copper and Si_3N_4 . Hence, the delamination has occurred either at the $\text{Cu}/\text{Si}_3\text{N}_4$ interface or at the $\text{Si}_3\text{N}_4/\text{SiO}_2$ interface. The nitride layer is 40nm thick and is used in order to avoid the copper oxidation during the process steps. Yet, due to the thickness of the nitride layer, conventional four point bending and cross section nanoindentation do not permit to get and compare these adhesion energies. Thus, it is not possible to discriminate these interfaces. On a few samples, a delamination can also be observed under the M4 layer, at the SiO_2/Cu interface, as depicted in Figure IV.29.

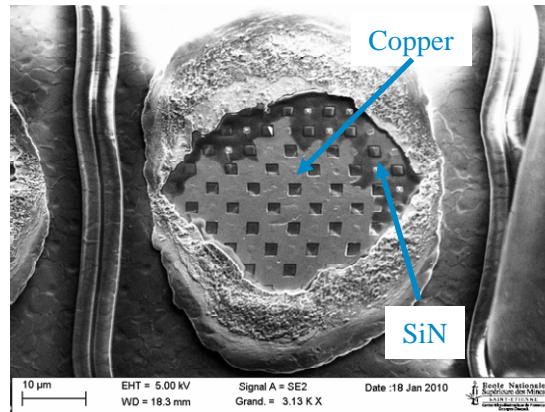


Figure IV.28. SEM top view of pad D: peeling is observed above M4.

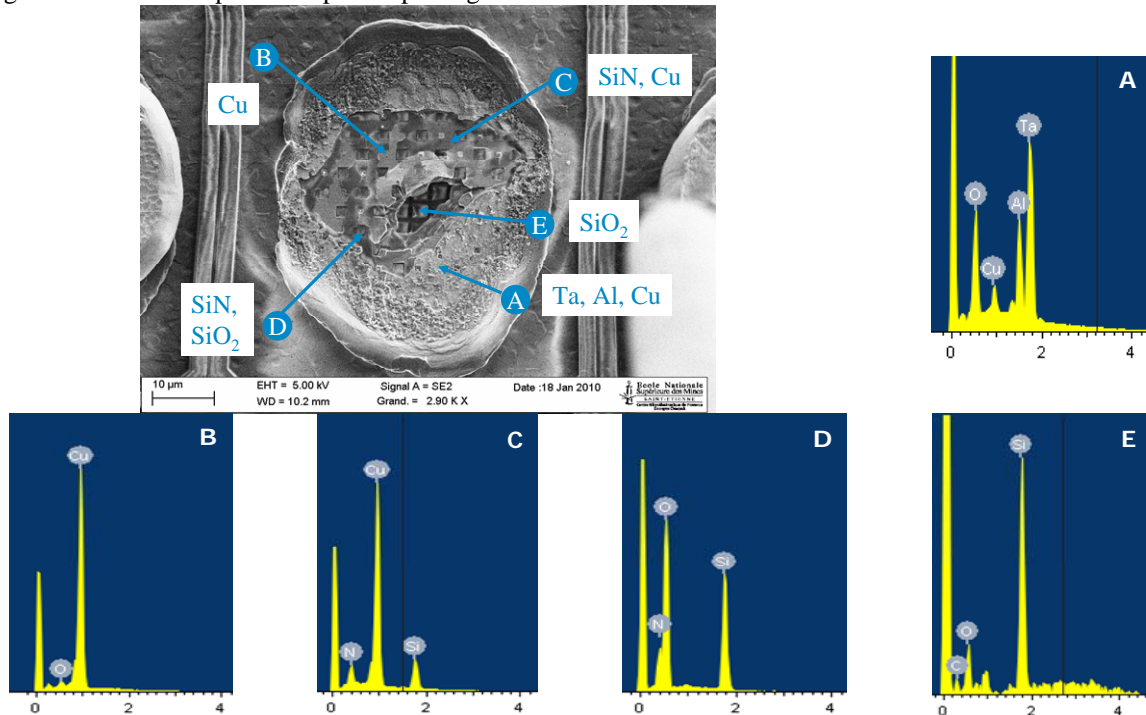


Figure IV.29. SEM top view of pad A, and corresponding EDX analyses.

On the pads without pad open layer (C, E, F and G), the delamination is necessarily different since the interfaces involved in the other structures do not exist in that case. The Figure IV.30 and Figure IV.31 show that the delamination occurs in two locations. On the one hand, just above the M3 layer, as depicted by EDX analyses in A on the 2 samples. The weak interface in that case is Cu/Si₃N₄. The M4 layer is not involved since it has another pad design (see for instance Figure IV.28), and the layer is torn off from the pad with some of it remaining in the outer of the peeled zone (see EDX analysis in location D on the Figure IV.31). On the other hand, another delaminated interface is observed between the oxide from the IMD2 layer and the M3 copper, as shown by the EDX analyses in the location C and E. Thus, for these pad structures, the peeling location is found to be either just above or just below the M3 copper layer. The path of the peeling generally follows the design of the M3 layer and the geometrical singularities between oxide and copper within M3 level, as depicted by the Figure IV.30.

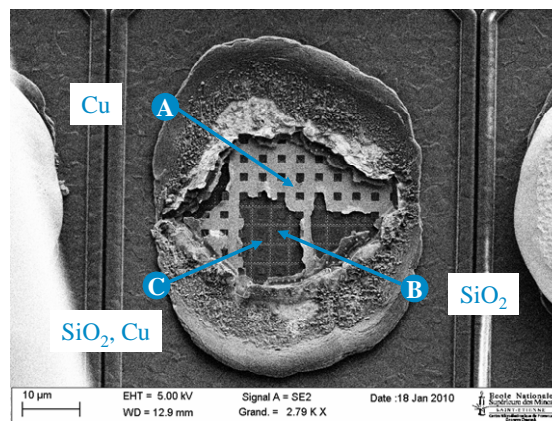


Figure IV.30. SEM top view of pad C: peeling is observed above and below M3.

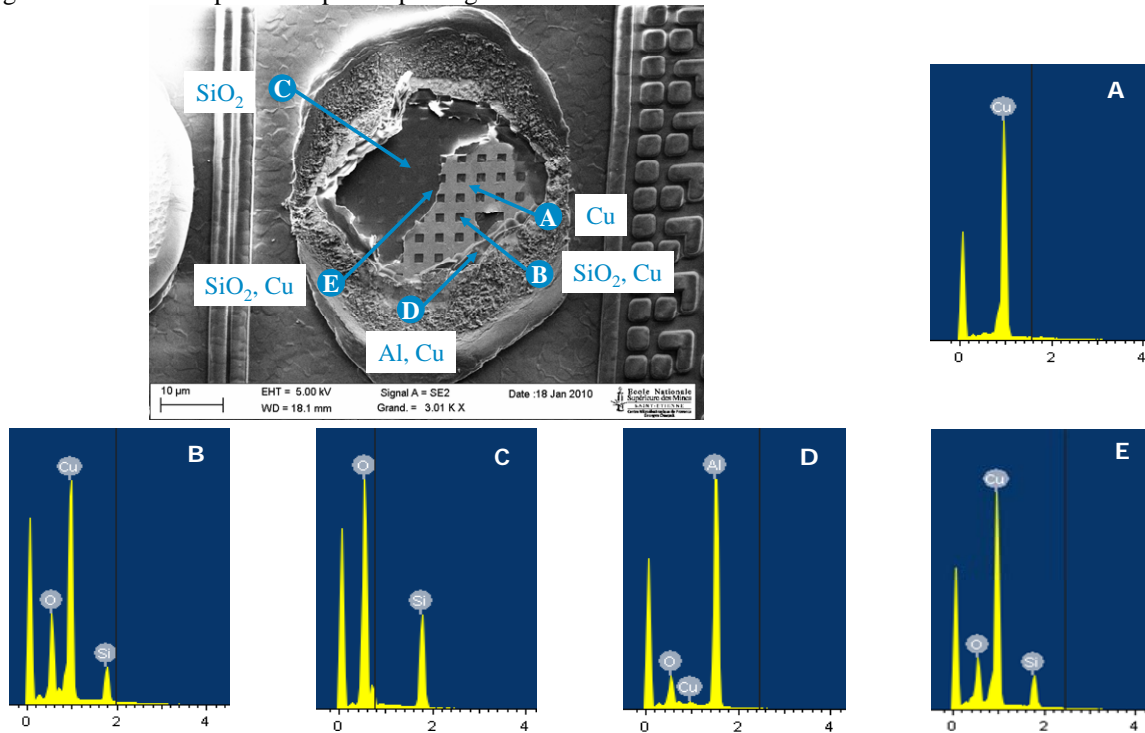


Figure IV.31. SEM top view of peeled pad G and related EDX analyses of the pad.

II.A.2.c) Discussion

During the wire pull tests, the metal density and the presence of vias have not been correlated with the peeling occurrence. In the current DOE, the copper density of the M4 is always 80%. Additionally, this layer is very thick compared to the other three metal layers (i.e. about 2 to 3 times thicker), which increases this effect. Hence, it underlines the supposed effect of the top copper layer for the pad resistance to peeling.

Concerning the peeling cases for the pad structures with pad open vias, the presence of the vias is related to several geometrical singularities, which have been shown to be the location of the failure initiation within the copper interconnections by Brillet-Rouxel and co-workers [Brillet-Rouxel, 2006]. The singular points for such layouts are shown in the Figure IV.32. Additionally, the top SEM views show that the crack initiate in these singular points, then propagates and reaches the surface at the edges of the ball. In the section III.D.2 in the previous chapter, the correlations between the process and the layout parameters have been shown. Therefore, a very strong interaction is found between the bonding power and the presence of these vias. The shear motion induced increases the stress in the structure, and particularly at the geometrical singularities. Similar trend and reasons are found while the contact force decreases.

For the pads without pad open oxide, the peeling is found just above or below M3, which is deeper into the structure (about $2.4\mu\text{m}$ or $2.7\mu\text{m}$ below the surface versus $1.8\mu\text{m}$ for pads with PO vias). The different ultrasonic (shear) stress repartitions and the induced residual stresses into the structures possibly explain these variations. The borders of the peeled areas rise. It emphasizes similar failure propagation from the M3 level to the surface of the pad, as depicted in the Figure IV.33.

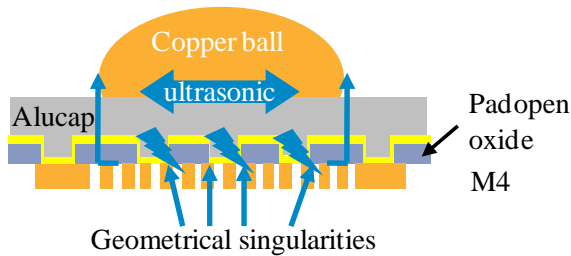


Figure IV.32. Schematic representation of the pad structures with PO vias, depicting the geometrical singularities.

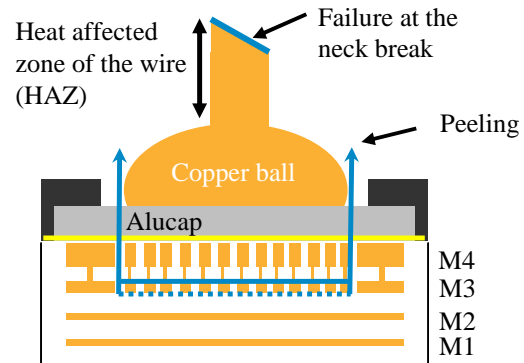


Figure IV.33. Schematic representation of the pad without pad open showing the peeling initiation and evolution.

II.A.3. Shear tests

The shear tests are made on 30 pads for each structure and bonding set of parameters to evaluate with quite good reproducibility the adhesion of the ball to the pad. Hence, 2160 pull tests are performed. This study follows the one started in section III.C.2 in the chapter 3. The failures mode 1 is observed (pad lift). The measured shear forces are presented in the Figure IV.34. Increasing the metal density augments the ball adhesion to the pad, which is related to a reduction of the overall stiffness of the pad structure and an improved compliance. Nevertheless, neither the presence of pad open vias nor the vias modifies the shear force.

Considering that the ball dimension is not affected by the layout parameters, the shear strength calculated from these measurements provides the same results.

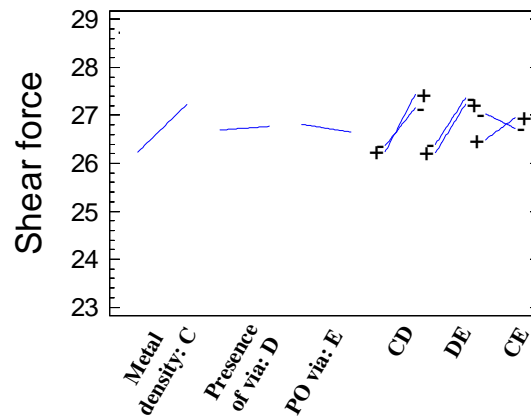


Figure IV.34. Pad layout effect on measured shear forces.

Hence, according to these results, the adhesion of the ball on the pad mainly depends on the process parameters. The design ones, which modify the overall stiffness of the structure, are able to modify the shear force only in small ranges if acting on the ball shear force.

II.A.4. Discussion

In this section, the results are faced to the ones found in the literature. The latter are summarized in the Table IV.1. For the numerical analyses, either a stress analysis or a released energy based criterion is used to establish a comparison between the designs. The energy released during the crack propagation is evaluated with the methods ARE or NRE: Area or Nodal release energy. For more details about the methodologies, refer to papers from van Gils et al. [van Gils, 2005] for the ARE method or Gallois-Garreignot et al [Gallois-Garreignot, 2007] for the NRE one.

Investigated parameters	Analysis method	Results and observation
Mercado et al. [Mercado, 2000]		
-Number of metal layers	Experiments and modeling	Experimentally, pad structure with 3 metal layers is more prone to ball shear and wire pull critical failure modes. Numerically, only minor differences are found in the relative stress concentration between 3 and 5 metal layers.
Chin et al. [Chin, 2007]		
-Number of vias	Numerical model (explicit)	The equivalent stress in the pad increases while decreasing the number of vias.
Hsu et al. [Hsu, 2007]		
-Number of vias -Via height -Via width (i.e. metal density) -Via distribution	Numerical model (similar to [Chin, 2007])	Equivalent stress distribution in USG increases while: -Reducing the number of vias, -Increasing the via height, -Reducing the via width, -Modifying the via distribution (loop vs. straight).
Kregting et al. [Kregting, 2006]		
-Peeling location -Metal lines -Dielectric layer -Via structure	Experiments and modeling	Metal peel off is initiated by delaminations under the alucap. The peeling location is affected by the design parameters from stress analysis and the use of ARE method.

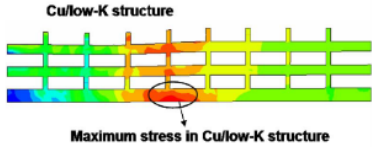
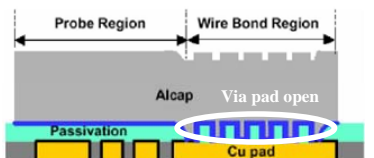
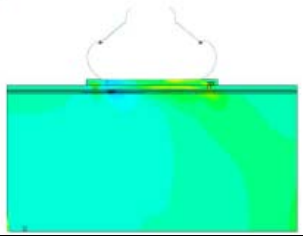
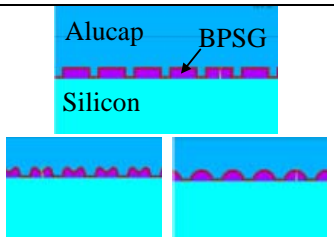
Engelen et al. [Engelen, 2007]		
-Metal density -Presence of via	Numerical model (ARE)	Mechanical performance of the whole stack is affected by the presence of via and the metal density.
Viswanath et al. [Viswanath, 2007]		
-Bond pad thickness -USG Young's modulus	Numerical model (explicit)	Von Mises stress increases while reducing the bond pad thickness. E_{USG} is not significant.
		
Lee et al. [Lee, 2007]		
-Presence of PO via	Experiments	Alucap lift is observed at the interface between Ta and Cu. The presence of PO via is required to correct this issue.
		
Liu et al. [Liu, 2008]		
-Bond pad thickness -Presence of TiW -Presence of vias	Numerical model (implicit)	Increasing the bond pad thickness induces a decrease of the stress in the structure. On the addition of TiW and vias, a trade-off is required.
		
Liu et al. [Liu, 2008]		
-Shape of the BPSG layer (dome shape vs. M shape vs. square shape) ¹	Experiments and Numerical model (similar to [Liu, 2008])	Neither crater in the silicon, nor stress variation is observed while modifying the BPSG shape.
		

Table IV.4. Design effects during wire bonding experiments and simulations found in the literature.

Hence, the studies evaluate many different design parameters and possibilities. Most of them are evaluated through numerical studies, and found to have an effect of the robustness of the pad. More precisely, an agreement is obtained concerning the increase of the pad thickness in order to reduce the peeling hazard. For the presence of via, results from Hsu and co-workers contradict those obtained by Liu et al. Lee et al. suggest to add PO via to avoid the delamination below the aluminium layer, however, in our study, these vias are shown to be detrimental to the pad. The metal density is shown to have an effect on the resistance to peeling by Engelen et al. and Hsu et al.

Several other studies evaluate the pad overall stiffness in the wire bonding context, whatever the bonding material. Hence, Degryse and co-workers [Degryse, 2004] showed numerically that the stress in the low-k becomes very high during the bonding process due to the reduced stiffness of the oxide material causing a large deformation. Similarly, by means of a stress analysis, Yeh et al. [Yeh, 2006] showed the effect of the material properties (Young's modulus and yield stress of the low-k, Young's modulus of the IMD layers) on the stress inside the pad structure. Experimentally, Ng et al. [Ng, 2006] found that using a harder ENIG

¹ In the structure modelled by Liu et al. [Liu, 2008], the BoroPhosphoSilicate Glass (BPSG) layer is located just below the aluminium layer. Then, modifying the shape of the BPSG can be closed to a modification of the shape of the pad open oxide vias.

(electroless nickel immersion gold) pad is related to larger ball diameter and an increase of the shear force.

Finally, the parameters evaluated during the wire bonding experiments are generally found to have an influence on the mechanical robustness of the pad. In our case, the effect of the very dense M4 layer is confirmed to reduce the perspectives of evaluating the effects of the metal density and the presence of via. Indeed, the M4 layer appears to have a shield effect for the evaluation of the layers below. Concerning the pad open via, differences in observation between the current study and others is probably due to other design effects such as the aluminium thickness of the design of these PO vias.

II.B. Nanoindentation tests on CMOS 120nm

II.B.1. Nanoindentation with Berkovich tip

On the pads used during wire bonding and associated characterizing tests, 15 indents are performed with the Berkovich tip.

The load-displacement curves for the pad E are shown in the Figure IV.35. A good reproducibility is obtained during the tests, whatever the pad structure. As seen in the previous case, no crack in the structure is obtained with this tip, and then the comparison between the curves is made using the maximum displacement into the surface, reached at the end of the loading. The results are depicted in the Figure IV.36.

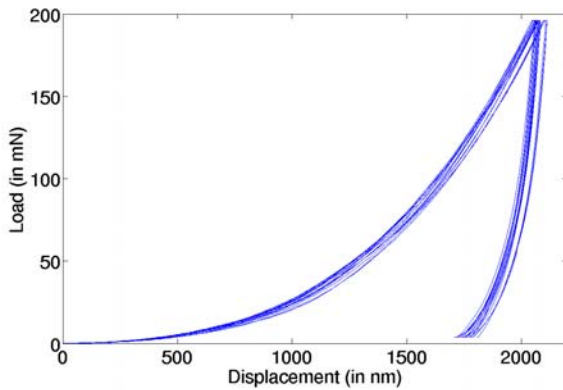


Figure IV.35. Load displacement curves for the pad structure E.

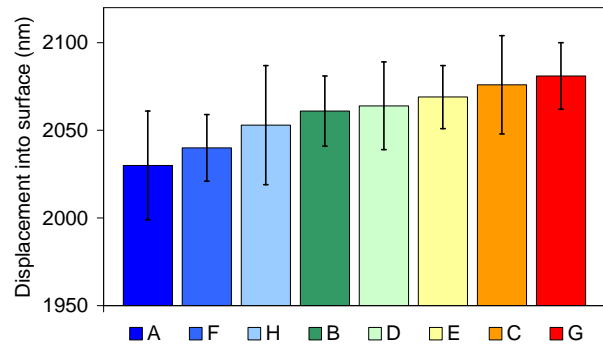


Figure IV.36. Maximum displacement into the surface for the various structures.

Low metal density and the presence of the pad open vias are seen to decrease the indent maximum depth. The difference between this trend and the peeling occurrence is further discussed.

The average displacements measured are very close. Yet, due to the very good tests reproducibility, a ranking according to the maximum displacements into the pad can be deduced:

$$G \approx C < E \approx D \approx B < H < F < A \quad (IV.6)$$

This ranking is submitted to uncertainty due to the small range of values (between 2030nm and 2080nm) and the closeness between the maximum depths. Thus, confirmation with the cube corner tip is needed, since the tip exhibits more local behaviours.

II.B.2. Nanoindentation with cube corner tip

Then, 15 indents are performed on the pads with the cube corner tip. The maximum load is 2gf or 19.6mN. The test is repeated twice due to reproducibility issues (series 1 and 2). The

top view of these indents is seen in the Figure IV.37. The view (b) shows that nanoindentation is necessarily made on the small vias in the oxide layer.

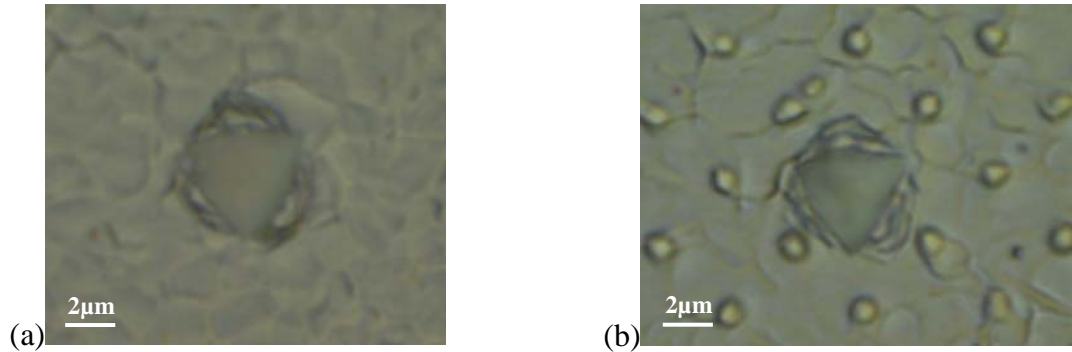


Figure IV.37. Optical microscopy top views of the pads without pad open via (a) and with PO vias (b), after nanoindentation.

The maximum displacement into the surface for all the structures is shown in the Figure IV.38. As mentioned, for some of the sets of indentations, which are the series 1 of structures B, D, F or E, and the series 2 for pad C, the reproducibility is very low. The structures with pad open via (B, H, A and D) are clearly differentiated to those without (F, E, C and G), and these layouts are split in the following discussion. The former pads have a lower maximum indentation depth than the latter.

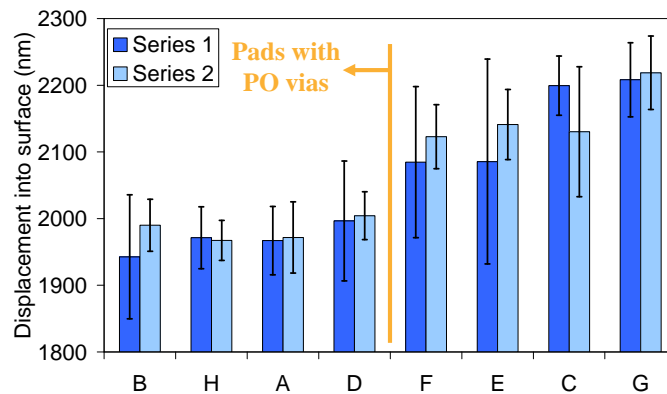


Figure IV.38. Displacements at maximum load for the various pad structures.

II.B.2.a) Pads with pad open vias

The pads with pad open vias are B, H, A and D. While going downward into the structure, the tip goes through aluminium, oxide and ends in the M4 layer. Hence, for these pads, the variability can come from surface detection due to the pad layout or from the penetration into the M4 layer, i.e. an effect of the indentation location on the pad.

The Figure IV.39 depicts the load-displacement curves for one of these structures, where cracks are clearly observed. These failures are analysed using the indicator described in the section I.B.2.c) and shown in Figure IV.40. For all the structures, the crack depth is found between 1450nm and 1500nm. As a consequence, they are located within the oxide layer. The crack load is found between 8.7mN and 9.3mN. Thus, the crack load and depth for the structures show little variation. This result is expected because only the metal density and the presence of vias are modified between all these structures, and these layers are located below the M4 layer. Similarly to the previous tests, a slope break is observed between the aluminium

and the oxide layer. Consequently, the local probing induced by the cube corner tip is confirmed.

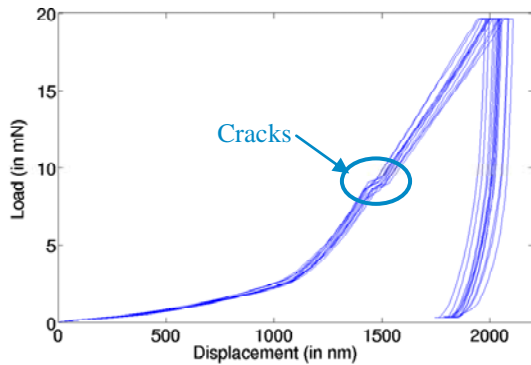


Figure IV.39. Load displacement curves measured for the test 2 of the pad D.

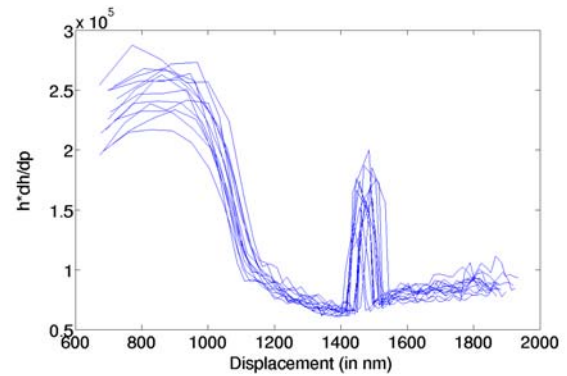


Figure IV.40. $h \cdot dh/dp$ curves as a function of the displacement for the corresponding measurements.

The SEM cross section views of the pad after indentation are showed in Figure IV.41. On the top view of the pad, a white area is observed in the centre of the indentation mark. This can be related to the tantalum layer observed, which means that during such test, the aluminium is completely removed in the centre of the indent, as demonstrated by the cross section in C. On the views A and B, a crack, which is observed on the measurements, is found in the pad open oxide layer. This failure is located at geometrical singularities due to the M4 design, which have been also pointed out during the wire pull tests.

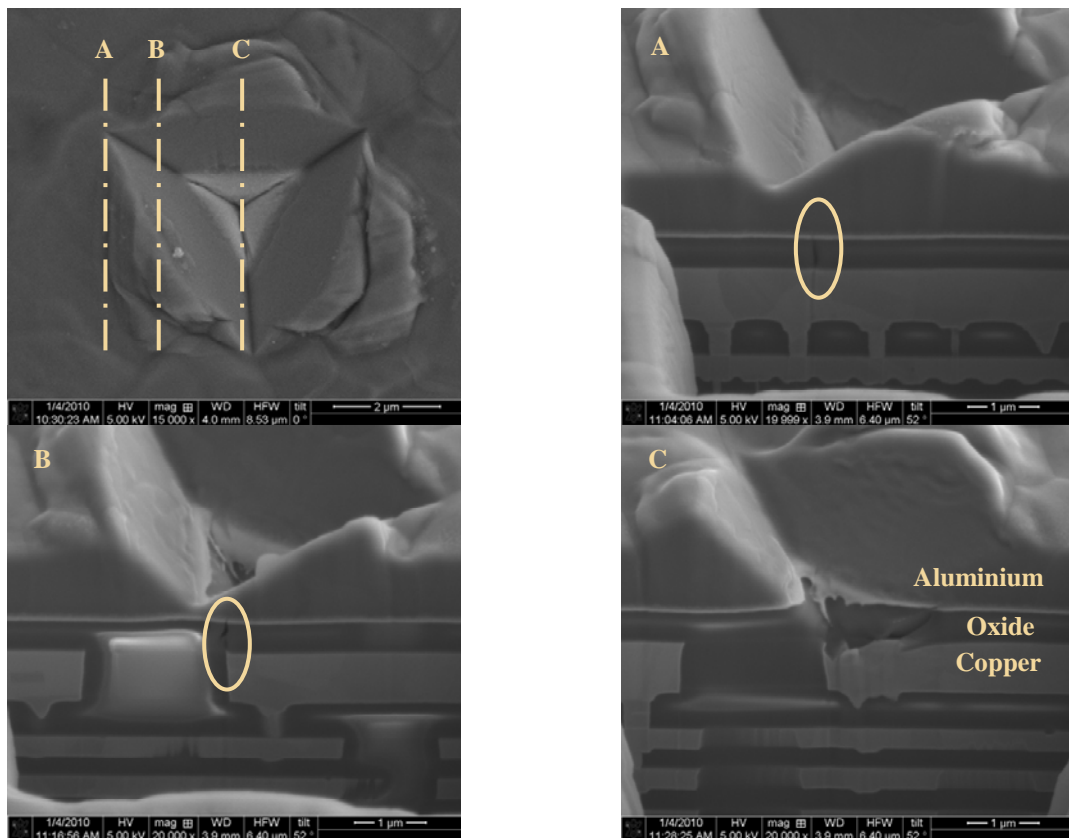


Figure IV.41. SEM cross section views after FIB milling of indented pad D, within crack is observed in the pad open oxide layer.

II.B.2.b) Pads without pad open vias

The pads without PO vias are D, E, C and G. While going downward into the structure, the tip goes through aluminium, and then reaches the M4 copper layer, which is composed of 80% copper.

The load-displacement curves corresponding to the pad G are depicted in the Figure IV.42. In spite of its increase along the loading stage, the slopes of the curves do not show a clear break at the interface between aluminium and copper. It is related to the material properties of these two materials. Indeed, the Young's modulus and hardness of the copper are respectively 132GPa and 1.1GPa from nanoindentation measurements¹. Hence, the variation of these properties (and particularly the E/H ratio) is lower than these between aluminium and oxide, and the effect reduced.

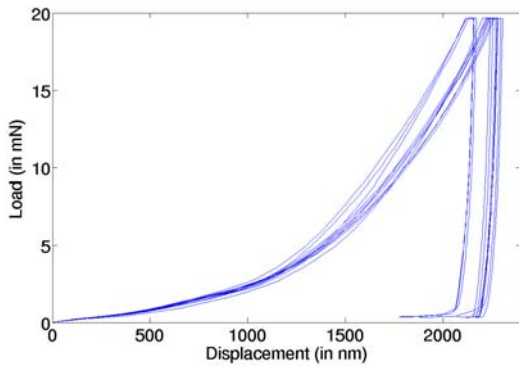


Figure IV.42. Load displacement curves measured for the test 1 of the pad G.

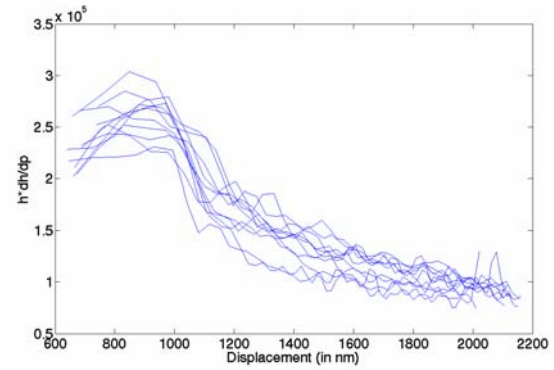


Figure IV.43. $h \cdot dh/dp$ curves as a function of the displacement for the corresponding measurements.

Moreover, no crack is observed on the measured curves, which is highlighted by the Figure IV.43, where no jump is observed on the curves. As a matter of fact, the tip penetrates through the copper layer, which is plastically deformed during the test, but is not fractured, as shown in the Figure IV.44. On the views, the centre of the indent is not as white as in case of nanoindentation in the pad with oxide beneath aluminium. Indeed, aluminium is found over the tantalum layer, even in the centre of the indent, as depicted in the Figure IV.44 b). This can be explained by the large plastic deformation of the copper.

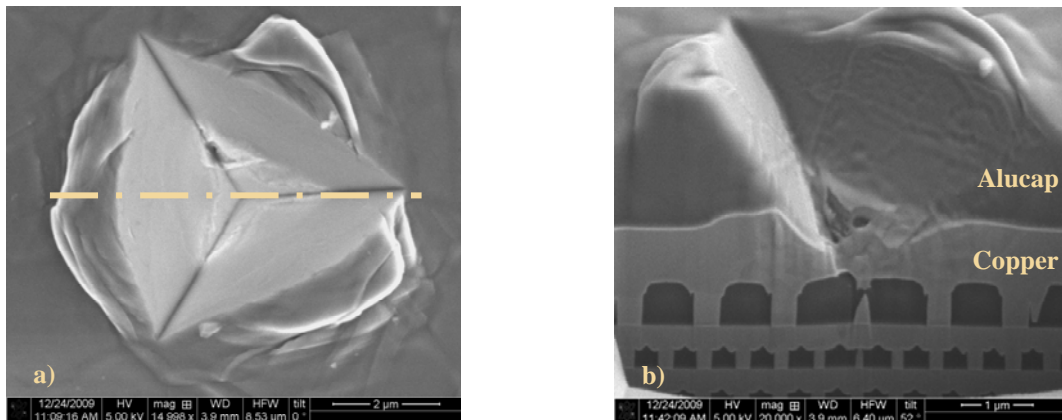


Figure IV.44. SEM top view of the pad (a) and SEM cross section view (b) after FIB milling on the indented pad.

¹ The measurement of the material properties of the copper is depicted in the chapter 5.

In these structures, the variation between the measurements is sometimes very important, such as measurements 1 on pads E and F and measurements 2 on pad C. It is induced by the M4 layer, which is made from 80% copper. Hence, in some particular cases, the nanoindentation takes place in the oxide filling the holes, or at the frontier between oxide and copper, which induces important variations. The load-displacement curves from the first measurement on the pad E are depicted on the Figure IV.45. Low reproducibility is found between the tests and the measurements on the samples 6 and 12 are more precisely observed. It is expected that the former penetrates in an aluminium over oxide stack and the latter an aluminium over copper one. The h^*dh/dp curves confirm this assumption for these two samples in the Figure IV.46. On the sample 6, a clear slope break and reduction of h^*dh/dp occurs at the interface between the aluminium and the oxide. Additionally, the peak corresponding to a small crack is highlighted in the PO. On the contrary, for the sample 12, no clear reduction or slope break is observed after indentation into aluminium, which characterizes an interface between aluminium and copper.

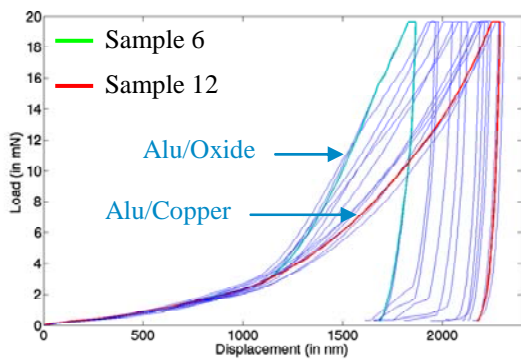


Figure IV.45. Measured load-displacement curves for the pad E, series 1.

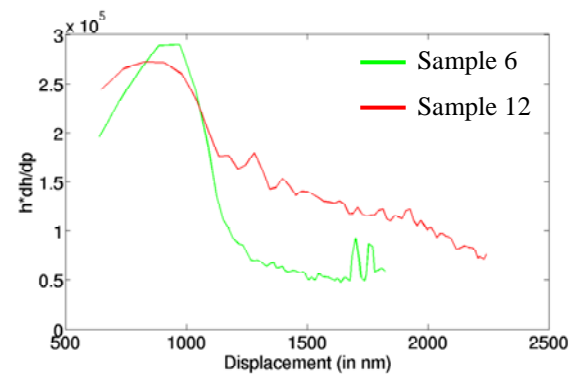


Figure IV.46. h^*dh/dp as a function of the displacement for the two samples investigated.

These results are confirmed by the SEM views following FIB milling along one of the median of the mark, for the two samples. On the Figure IV.47, for the sample 6, the indent penetrates through the aluminium to end in the oxide film. Yet, the crack is not detected during these milling and observations. Inversely, for the sample 12 shown on the Figure IV.48, the indentation evaluates aluminium over copper stack, even though little oxide part is also probed during the test.

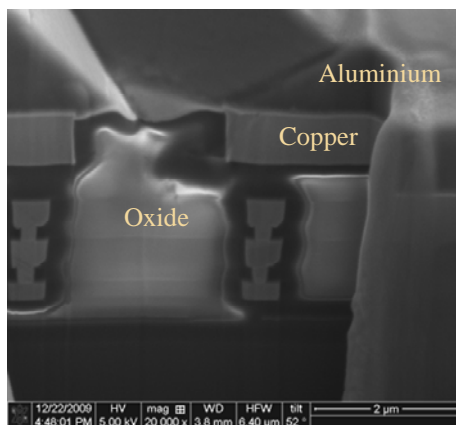


Figure IV.47. SEM cross section after FIB milling of the sample 6.

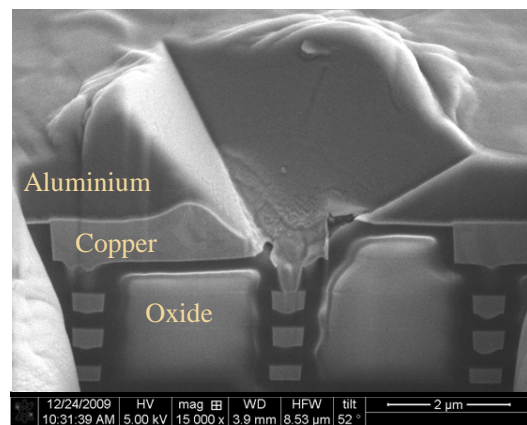


Figure IV.48. SEM cross section after FIB milling of the sample 12.

These tests and views underline the ability of nanoindentation to evaluate and detect very accurately the structures. The sensibility of the analysis while probing a pad with the cube corner tip is also demonstrated. However, this advantage can become a limitation due to the size and the sharpness of the cube corner tip that can probe very small details and local microstructure and not provide a global response of the structure robustness.

II.B.2.c) Ranking and discussion on the robustness of the pads

Considering the few uncertainties in the measurement reproduction aforementioned, it is possible to weight the structures according to their mechanical robustness, taking into account the average values for each stack, over the two tests.

Thus, the ranking according to the maximum displacement into the pad (Figure IV.38) is:

$$G < C < E \leq F \ll D < H < A < B \quad (\text{IV.7})$$

Inserting pad open via in the structure is obviously reducing the maximum penetration depth, due to the difference of material properties between oxide and copper. This effect is clearly observed during this set of measurements.

The decrease of the metal density from 80% to 30% induces a little reduction of the maximum indentation depth. The presence of via is not found to have a significant effect on the displacement reached at maximum load. These results confirm the local testing achieved during nanoindentation with cube corner tip and the high resistance to compression of the oxide faced to copper.

II.B.3. Discussion: wire bonding versus nanoindentation

During the wire pull and shear tests, the effect of the pad open vias has been clearly emphasized as detrimental to the structure. Additionally, the metal density and the presence of vias showed no clear effect on the pad robustness. The following ranking according to the peeling rate has been obtained (IV.5) for the pad structures:

$$D < H \ll A < B \ll C < F \approx G < E$$

The bonding operation is related to a compression applied to the pad through the contact force and then the bonding force, and to a shear due to the application of the ultrasonic power. The wire pull test induces a tensile load to the structure. This load is not perfectly normal to the pad since the hook cannot be located right above it. Hence, the pull angle should also be considered while interpreting the pull tests results, as shown by Fiori et al. [Fiori, 2007].

The nanoindentation tests consist mainly in a compressive load applied to the pad¹. Thus, the loading undergone by the pad is much more complex in case of peeling. The depth reached at maximum load with Berkovich and cube corner tips depict the same trends, but both submitted to some uncertainties due to reasons described in the previous paragraphs. The ranking according to the depth is for the Berkovich and cube corner tip respectively:

$$G \approx C < E < D \approx B < H < F < A$$

$$G < C < E \leq F \ll D < H < A < B$$

Hence, the classifications are quite similar. The main difference is that Berkovich tip underlines the effect of the metal density, while the cube corner one this of the pad open oxide layer. This difference is explained by the shapes of the two tips, which probe more the volume or the local properties of the pad. Both studies show that the presence of PO vias, and at a lower degree the reduction of the metal density increase the pad resistance to compressive loading.

¹ At the vicinity of the indenter tip, some shear forces are also applied to the pad due to the resulting forces and the tip geometry.

The peeling rate demonstrates that the use of pad open via is detrimental to the structure. On the contrary, the nanoindentation tests have shown that the oxide layer between the aluminium and the copper increases the resistance of the pad to compression. As a consequence, on that particular sample, it seems that the compressive loading is not directly held responsible of the peeling observed. The compressive load and more particularly the reduction of the contact force has been observed in the previous chapter to induce peeling. The assumption of the reduced surface application of the ultrasonic power can be related to that result. In a more general point of view, the presence of oxide between aluminium and copper is shown to increase the resistance to compression. Therefore, the structure with only M4 below aluminium implies lowered resistance to compression, but no crack is observed inside the structure. A critical load prior to failure can also explain the difference observed between the results.

In appendix A, the effect of similar study with reduction of the maximum displacement into the surface due to the presence of a thick oxide layer beneath aluminium is shown. Actual probing and wire pull tests are performed and show the same ranking as the one deduced from the nanoindentation measurements. It underlines the effect of a critical load prior to failure and the negative effect of the pad open vias.

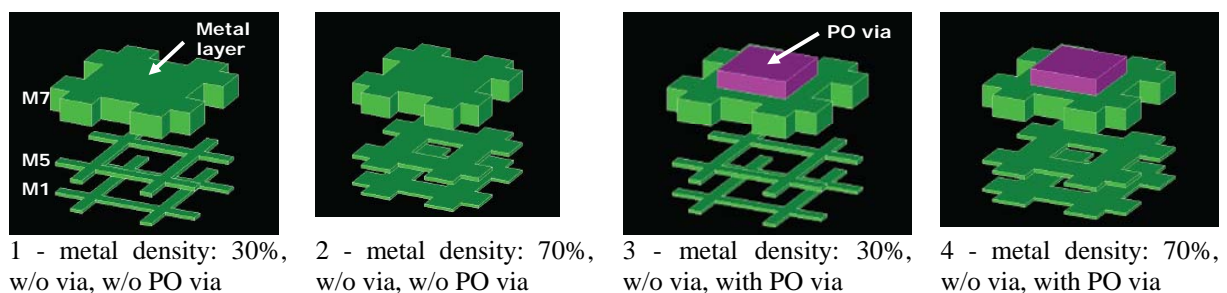
As a conclusion the nanoindentation with cube corner tip is generally able to evaluate the effect of pad open vias. However, for some cases the strengthening of the structure (due to the presence of oxide) is correlated to cracks, and critical load of the structures seems hard to determine.

II.C. Nanoindentation on CMOS 45 nm pads

As part of the calibration on EWS numerical model for the CMOS 45nm, these pads, with various structures have also been faced to nanoindentation.

II.C.1. Pad description

For this technology node, seven layers of metal are required to perform the interconnections. The thicknesses of the aluminium, pad open oxide, copper M7 and inter-metal dielectric layers composing the pad are respectively 1450nm, about 600nm, 900nm and 600nm. The structures investigated have distinct metal densities, presence of via and presence of pad open via. The metal density of only the five layers (M1 to M5) in the bottom of the structure is either 30% or 70%. The presence of via is investigated between all the copper layers, from M1 to M7. Contrarily to the tests on the CMOS 120nm, the option without via in the padopen layer consists of a full plate of oxide. The layouts are depicted in the Figure IV.49. It should be noticed that for lower levels (below M5), in this technology node, a low-k oxide is used instead of silicon glass. As mentioned on the first chapter, the mechanical properties of these layers are much weaker than these of TEOS. In their work, Wu et al. [Wu, 2006] have measured the film's mechanical properties with the Berkovich and the cube corner tips. For the Young's modulus and the hardness, they obtained respectively about 6.5GPa and 1.1GPa.



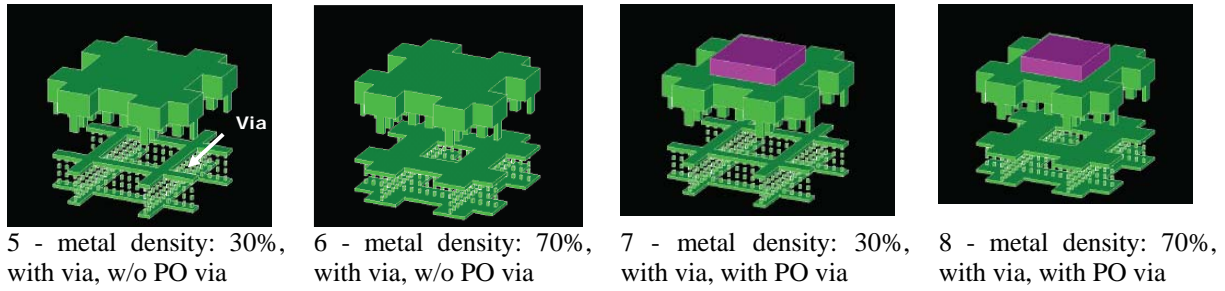


Figure IV.49. Various pad layouts evaluated during the nanoindentation on CMOS 45nm.

II.C.2. Measurements with Berkovich tip

Due to the reduced number of pads available on a die, 8 indents are performed for each sample and the maximum load is 20gf (196mN). Yet, the variability between the dies is checked, and for the pad layouts 1, 2, 5 and 6, the averages difference is below 7nm. The size of the indent after the tests is about $16\mu\text{m}$ for one triangle height of the projected area. Like the other nanoindentation tests, aluminium pile up is observed on the top view of the pad (Figure IV.54).

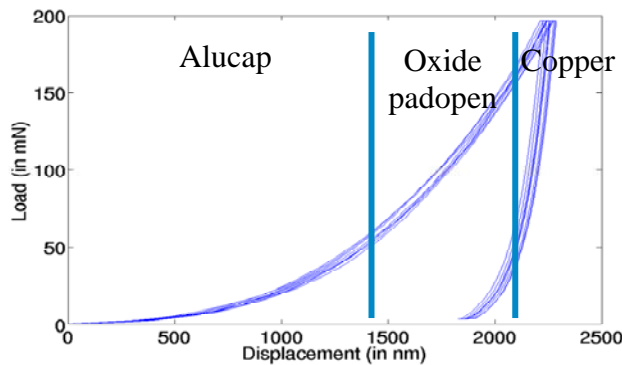


Figure IV.50. Load-displacement measurement for the pad layout 1.

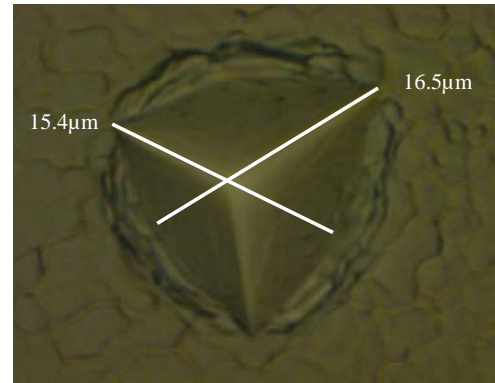


Figure IV.51. Optical microscopy top view of the pad after indentation.

The standard deviation is found below 27nm, which underlines the good reproducibility of the tests. The maximum displacement into surface for each pad design is given in the Table IV.5. For all the measurements, the depth reached at maximum load is found between 2194nm and 2303nm.

Pad layout	Metal density (%)	Presence of via	Presence of PO via	Maximum depth into surface and <i>standard deviation</i> (nm)
1	30	No	No	2248 (21)
2	70	No	No	2204 (7)
3	30	No	Yes	2303 (13)
4	70	No	Yes	2266 (27)
5	30	Yes	No	2234 (15)
6	70	Yes	No	2194 (18)
7	30	Yes	Yes	2294 (15)
8	70	Yes	Yes	2255 (21)

Table IV.5. Depth reached at maximum load and standard deviation for the various pad layouts.

The effect of the metal density of the lower layers is depicted in the Figure IV.52. The average and also the position of the minimum and maximum measured curves show that increasing the metal density from 30% to 70% is related to a decrease of the depth reached by

the indenter. This is observed for all the layouts, while other parameters are kept constant. The reduction of the depth is often about 45nm, which is higher than the standard deviation of the measurements. Hence, for structures with low-k, the increase of the metal density augments the resistance of the structure to compression. It can be directly related to the mechanical properties of the two materials.

The presence of pad open via is presented in the Figure IV.53. Their presence induces an increase of the displacement into the pad. The augmentation is about 55nm whatever of metal density and the presence of via. Considering the size of the indent, during the test, the oxide and the aluminium are probed, and an increase of the displacement into the surface is expected due to the presence of aluminium filling the vias (similarly to Figure IV.26 b)).

The presence of via reduces slightly the displacement into surface (about 10nm) for each structure compared. However, considering the standard deviation of the results, this variation is not significant. This can be explained since the addition of via into the structure induces a very low variation of the density of the layer where they are located due to their size and sparseness. However, the presence of via has to be related to the increase of the copper density with respect to the properties of the materials introduced.

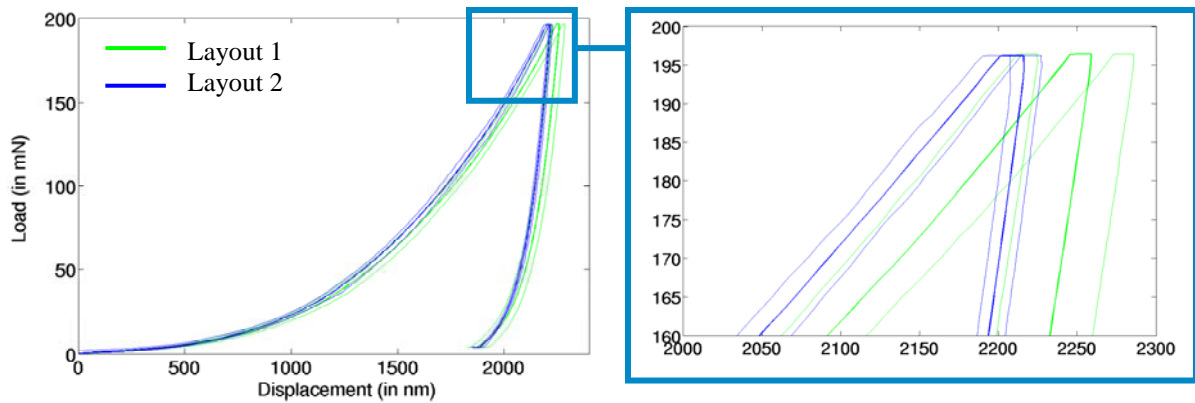


Figure IV.52. Minimum, average and maximum indentation curves for the pad layouts 1 and 2, depicting the effect of the metal density.

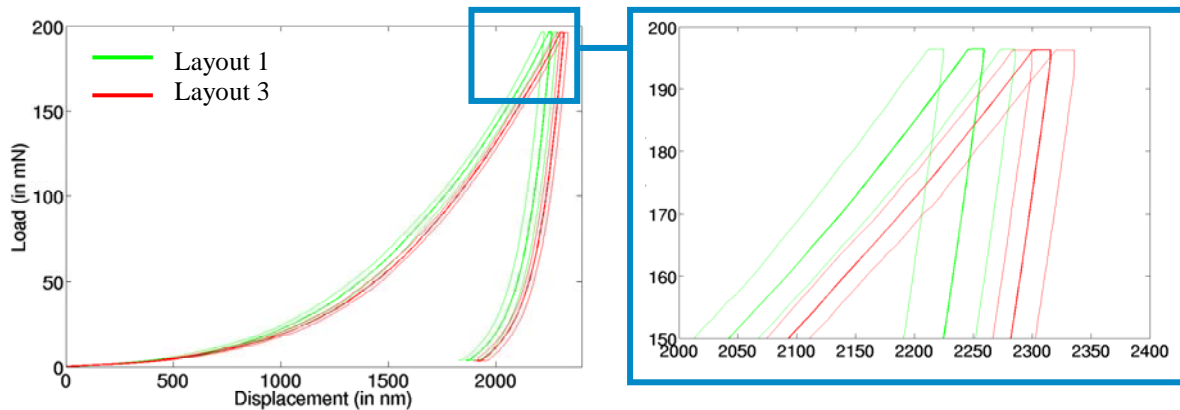


Figure IV.53. Minimum, average and maximum indentation curves for the pad layouts 1 and 3, depicting the effect of presence of pad open via.

The ranking according to the maximum displacement into the surface reached by the indenter at maximum load is:

$$3 \leq 7 < 4 \leq 8 \leq 1 \leq 5 < 2 \leq 6 \quad (\text{IV. 8})$$

II.D. Discussion on the pad layouts effect

The wire pull tests and the nanoindentation are used to evaluate the mechanical robustness of the structures. More particularly, the copper density, the presence of via and of pad open vias is questioned. The peeling occurrence have clearly underlined that the presence of the PO via under the bonding area is detrimental to the pad due to the presence of the geometrical singularities. The nanoindentation with both Berkovich and cube corner tips of the same structures have shown an opposite trend. It has been related to the difference of loading condition of both tests. The indentation tests made on CMOS 45nm structures have shown that the use of PO via induces a reduction of the resistance of the pad to compression compared to an oxide full sheet. Hence, the presence of padopen via is worse than the aluminium over copper stack and also the aluminium over oxide full sheet one.

The metal density of the copper layers has been investigated only for the lower levels for both technology nodes. The peeling rates have highlighted that the metal density have reduced effect, for the CMOS 120nm node. It has been related to a very thick and dense top copper layer that reduces the effect of the layers below on the overall resistance of the structure. For the nanoindentation test with Berkovich tip, the metal density effect has been observed. On the CMOS 120nm structure, increasing the metal density reduces the resistance to compression, and for the CMOS 45nm, the opposite results are observed. The difference is due to the presence of low-k oxide in the latter structure. The silicon glass has higher resistance to compression than the copper, but the low-k oxide has a lower one. Yet, the metal density is not necessarily directly related to issues in the pad. Therefore, for latest technology nodes, within low-k oxides are used between lower metal levels, increasing the copper density can improve the pad resistance to compression.

The presence of via, whatever the technology node has not been found to have an effect on the mechanical robustness of the pads during the current experimental tests. In the literature, Cecchetto et al. [Cecchetto, 2007] investigated the effects of the presence of via and the via pitch (1 μ m vs. 4 μ m) in the M3 layer (out of 5 metal layers, whose top one is 3 μ m thick aluminium) using a numerical model, during probing tests. The structure without vias in the M3 layer has been related to lower Von Mises stress and then, wider spacing between the vias is also shown to reduce the stress. Thus, this numerical results seems to be different from the current ones, however, higher stress is not always directly related to higher failure rates, if critical values are not obtained, which can explain the difference.

III. Metal copper layer design

In the previous sections, it has been observed that the failures in the pad open layer have a strong correlation with the layout of the copper layer below since their initiation starts at the copper geometrical singularities. As a consequence, the effect of the top copper layer's design is investigated (M4). The layout modification can induce small variation of the copper density. However, the densities of each of the layer are very close. The DOE performed to study the effect of EWS described in section I is split. The results for the structures made from long lines have been presented. Here, pad designs with slots are also considered. Once more, the comparison between the layouts is made from actual probing, nanoindentation with both tips and modelling.

III.A. Pad layouts description

The pad layouts investigated are of two kinds: lines and slots. The former correspond to long copper lines, separated with oxide, as shown in the Figure IV.54 a). The metal density

for this layout is 89%. For the latter, the copper is deposited in a way that oxide forms small islands. This layout is commonly known as slotted copper plates or slots. The copper density for such designs is 74%. This design is derived in three variants. Hence, pad layout with $1\mu\text{m}$ slots, $1\mu\text{m}$ slots with shifted plates and $2\mu\text{m}$ slots, as depicted in the Figure IV.54 b) to d). The layout with shifted copper plates is investigated to evaluate the effect of the location of oxide small line between the plates. In the standard design, it is located in the centre of the pad, thus, surely under the probing area. The use of larger slots can have an effect on the copper deposition, grain formation and the number of singularities. As a consequence, the overall mechanical robustness of the structure can be affected.

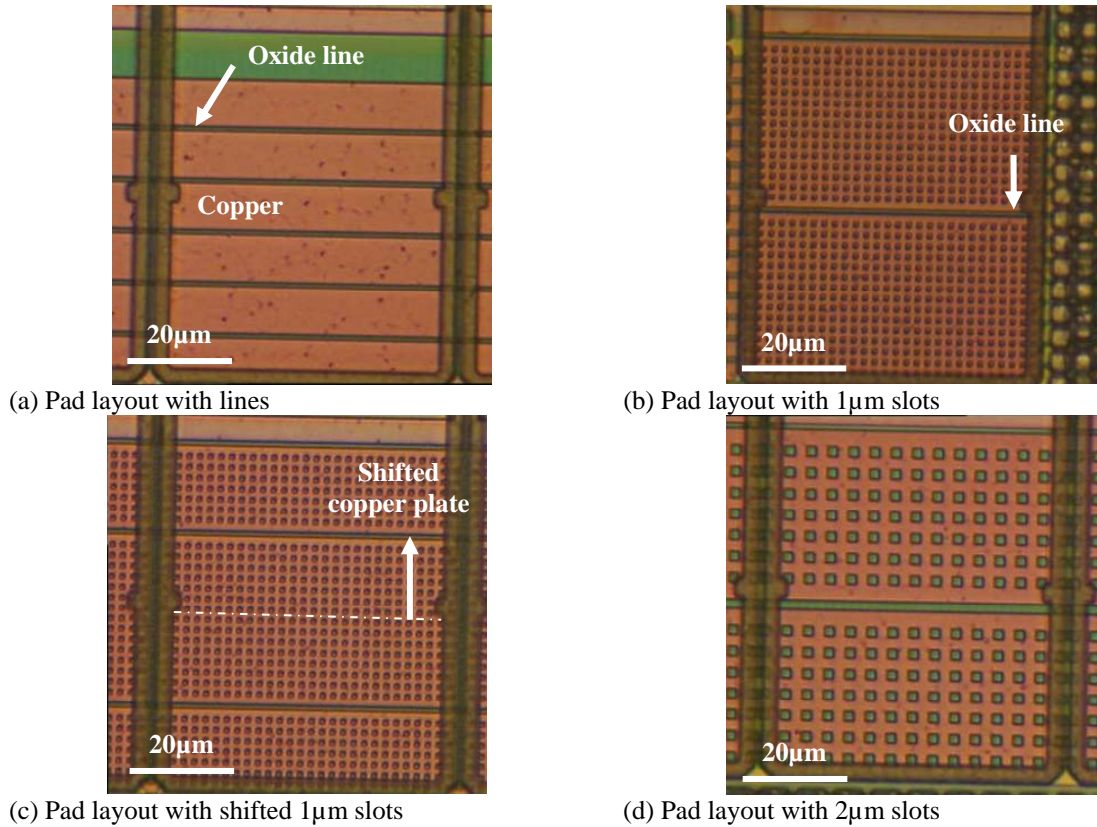


Figure IV.54. Optical microscopy top views of the pads after aluminium and tantalum removals.

III.B. Pad integrity tests

The EWS is performed on the pads with the harshest condition (OD $90\mu\text{m}$ and 10 passes). Then, similarly to other studies, the aluminium and tantalum layers are removed, and cracks are numbered. The pad integrity test is made on 4 dies containing 28 pads for lines structures and 40 pads for layouts with slots, thus, on respectively 112 and 160 pads. All the cracks that can be observed are located in the PO layer.

The results of the pad integrity tests are shown in the Figure IV.55. Whatever the stack, it is found that the highest crack occurrence is observed for the pad layout with lines. For the thinnest stack, the failure rate reaches 91.7% of the pads, which show major issues. Then, increasing the size of the slots is shown to increase the crack occurrence in the considered range of values. The most robust structure is the pad layout with $1\mu\text{m}$ slots on shifted plates, which is shown to enhance the mechanical robustness. Hence the position of the oxide lines between the copper plates has an effect on the crack occurrence.

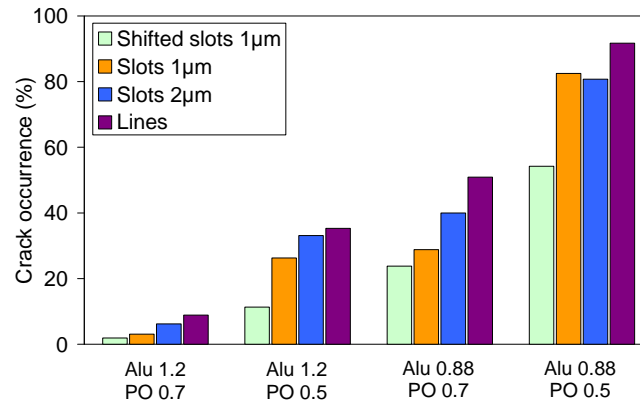
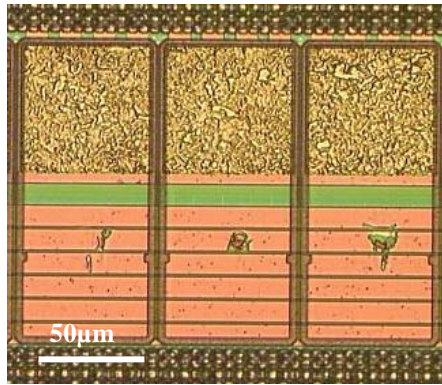
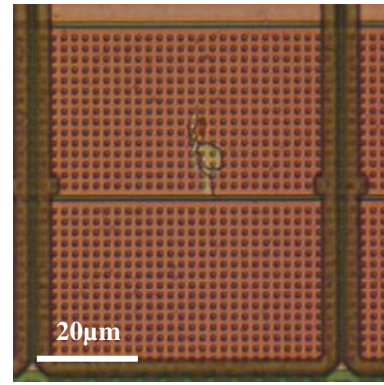


Figure IV.55. Crack occurrence for the various pad designs.

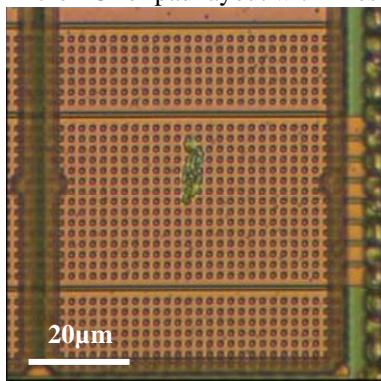
This typical crack for the structures is depicted in the Figure IV.56. On layouts (a), (b) and (d), the cracks propagate from or until the small oxide line between the lines or the plates. For the shifted copper plates (c), the crack propagates between the singularities related to the presence of the slots, but do not reach the oxide line.



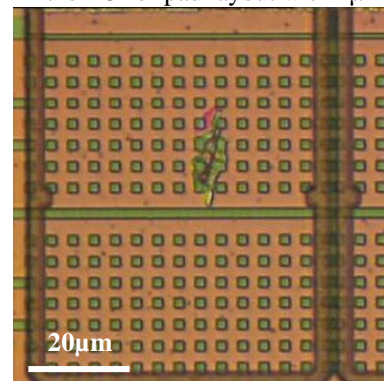
(a) Crack in the PO for pad layout with lines



(b) Crack in the PO for pad layout with 1μm slots



(c) Crack in the PO for pad layout with 1μm slots on shifted plates



(d) Crack in the PO for pad layout with 2μm slots

Figure IV.56. Optical microscopy top views of the pads after pad integrity test, depicting cracks in the pad open oxide layer.

Finally, based on the crack occurrence, the ranking of the layouts according to their mechanical robustness is:

$$\text{Lines} < 2\mu\text{m Slots} < 1\mu\text{m Slots} < 1\mu\text{m Slots with shifted plates} \quad (\text{IV. 9})$$

This provides the basis design ranking for the following comparison with the results of the nanoindentation tests and the numerical model.

III.C. Nanoindentation

In a second step, the nanoindentation apparatus is used to evaluate the influence of the layout on the robustness of the stack, using the Berkovich and then the cube corner tip.

III.C.1. Berkovich tip

The nanoindentation with the Berkovich tip is performed on 15 pads for each pad structure. The maximum displacement into the surface is collected in the Figure IV.57. The structure with lines shows higher maximum depth for all the stacks. The density of this layout combined with the remarks on the ability of oxide to resist to compression is one of the reasons. The plastic deformation of the copper lines shown in the Figure IV.18 also explains the current results. Due to the presence of small oxide cubes in the copper, harder and more brittle material, the M4 metal layer with slots has lower capability to plastically deform and bend under compressive loading. Accordingly, the pad layout with slots is related to higher resistance and thus lower displacement into the surface.

All the structures with slots are in a very close range. Yet, from comparison between all the stacks and considering the range of uncertainty between the measurements, the layout with 1µm slots provides the lowest average and related range of values, then the 1µm slots with shifted plates, and finally the 2µm slots.

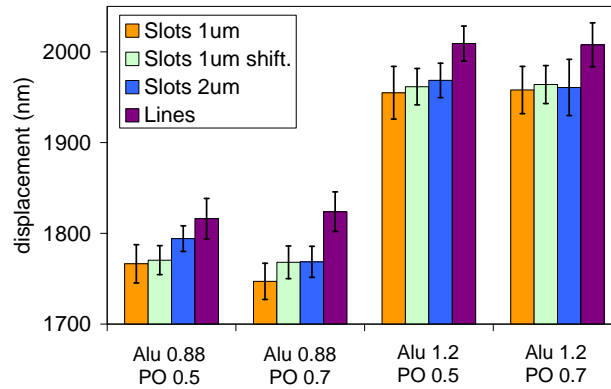


Figure IV.57. Maximum displacement into surface for the various pad layouts.

Hence, the ranking of the layouts according to their resistance to compression with the Berkovich tip is:

$$\text{Lines} \ll 2\mu\text{m Slots} \leq 1\mu\text{m Slots with shifted plates} \leq 1\mu\text{m Slots} \quad (\text{IV. 10})$$

III.C.2. Cube corner tip

In order to obtain good test reproducibility, 24 measurements are performed for each pad structure. The maximum indentation depth for all the structures is shown in the Figure IV.58. Whatever the thicknesses of the aluminium and pad open oxide films, the M4 design with lines is always the structure within the displacement into the pad is the highest. For pads with slots, the design with 2µm is related to the worst case. The two other ones show very similar results, with measurements a little lower for standard 1µm pads.

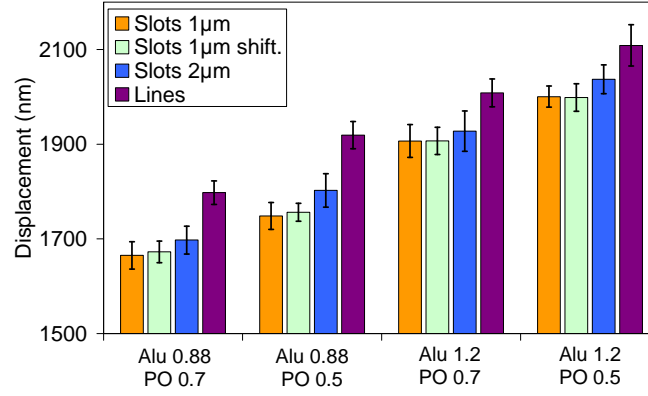


Figure IV.58. Maximum indentation depth for the various pad layout options and stacks.

On the nanoindentation curves, failures are observed on the load-displacement curves and their corresponding load and depth are extracted following the method aforementioned. In three out of the four stacks, the M4 design with lines is related to lower loading prior to failure, as depicted in the Figure IV.59. Nevertheless, the difference is very low compared to the uncertainty and the other process and stacking effects. All the structures have roughly the same load and displacement for appearance of the cracks. In all the cases, the displacement corresponding to the fracture is close to half the thickness of the oxide film. The location of the failure does not depend on the thickness of the oxide film. It is emphasized by the about 100nm difference of crack depth between the stacks, whose oxide thicknesses are 0.5µm and 0.7µm.

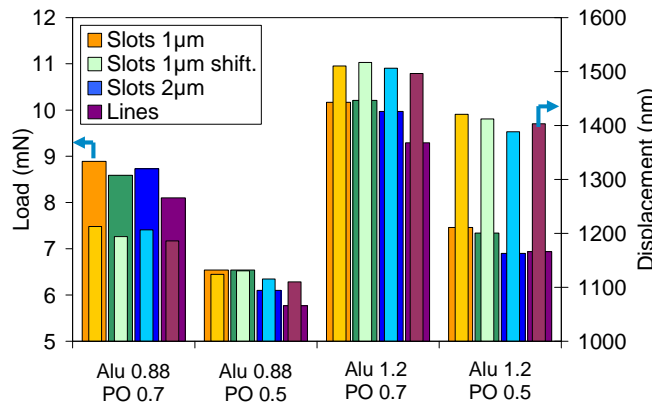


Figure IV.59. Load and depth of the crack for the various stacks and layouts.

To summarize, the ranking according to all the indicators described is:

$$\text{Lines} \ll 2\mu\text{m Slots} < 1\mu\text{m Slots with shifted plates} \leq 1\mu\text{m Slots} \quad (\text{IV. 11})$$

The effect of the pad layout is lower than the thickness one, and mechanical robustness can be modified in a smaller range.

III.D. Numerical results

The explicit finite element model described in the chapter 2 is used to compare the various pad layouts. For this analysis, the small details cannot be introduced in the numerical model, because it would lead to a very high CPU time consumption due to a high number of elements in the model. Hence, the M4 layer is homogenized, using the technique described in the second chapter. The material properties resulting have also been given.

For the pad layout with slots, the construction is made in a way that the structure with $2\mu\text{m}$ slots is a homothecy with ratio equal to two. The pad layout with shifted plates is made on the basis of the standard one. As a consequence, the three pad layouts with slots have the same design along the 3 axes and copper density in the M4 film. Thus, the homogenization technique lead to the same elastic orthotropic material properties and these layouts cannot be distinguished, leading to the same results. The use of the orthotropic properties is necessary due to the differences of copper densities in the two directions of the layer's plane. Then, only the comparison between the layouts with slots and with lines is achieved.

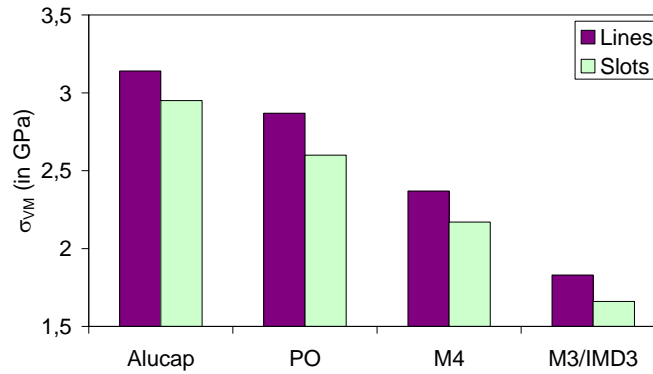


Figure IV.60. Maximum Von Mises stress in the layers for the pad with highest thicknesses of aluminium and pad open oxide.

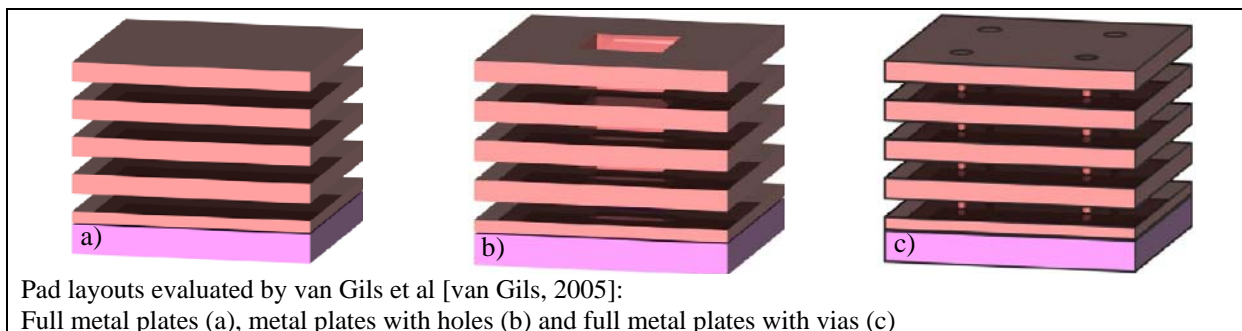
The analysis based on the Von Mises stress in the top layers of the pad is shown in the Figure IV.60. The stress in the oxide layers (PO and M3/IMD3) decreases while designing pad with slots. Hence, introducing slots with lower copper density, which also lead to a reduction of the equivalent Young's modulus, shear modulus and density of the layer, induces a reduction of the stress.

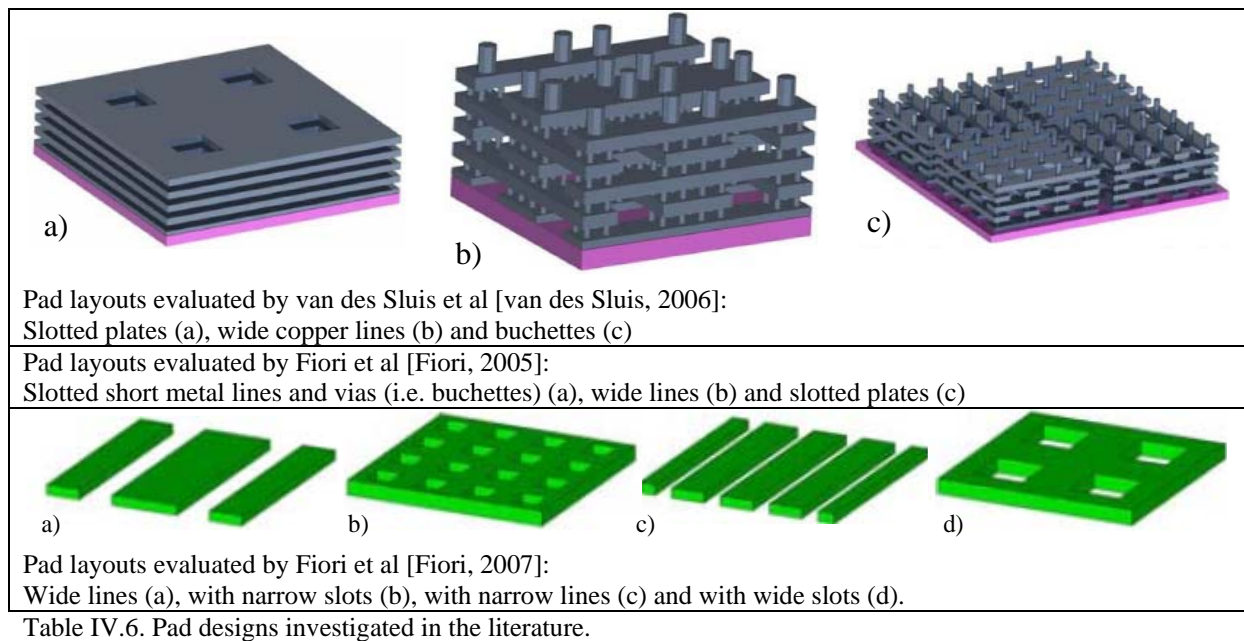
Finally, the ranking according to the Von Mises stress analysis for the pad is:

$$\text{Lines} < \text{Slots} \quad (\text{IV. 12})$$

III.E. Discussion

The design of the pad is not often investigated in the literature, and some of the structures evaluated are depicted in the Table IV.6. Even though their main application is the resistance to the wire bonding process and tests, since part of the pads have to resist to EWS and wire bonding, trade-off and understanding of the issues induced by both is required. All the studies are made on results from numerical models, sometimes confirmed by actual tests.





In their paper, van Gils and co-workers [van Gils, 2005] compared the various layouts using the area release energy method. A positive effect of the presence of holes (slotted plates) and of the presence of via is observed. Also using the ARE method, van der Sluis et al. [van der Sluis, 2006] found that the structure b) has a released energy slightly below the one of the layout a) and structure c) is the worst one. These results were confirmed by experimental tests. The absence of via in the structure a) while comparing it to b) may explain the differences found in the study and in those previously demonstrated.

In their work, Fiori et al. [Fiori, 2005] also obtained that the first principal stress for layout with buchettes is higher than this of the other structures. The numerical analysis, similarly to the experimental tests, was not able to distinct the structures with wide lines and slotted plates. In a more recent paper, Fiori and co-workers [Fiori, 2007] compared pad layouts with lines and slots and each of them has been derived with two distinct sizes. Indeed $3\mu\text{m}$ and $6\mu\text{m}$ oxide holes in the plates are investigated using the NRE method. Then, the ranking between the structures is established. The layout with wide slots (d) is the best one, followed by the layout with narrow slots, the layout with wide lines and finally the layout with narrow lines. An experimental DOE is performed, and a ranking according to the peeling rate confirmed the numerical results. Yet, the structure d) has not been experimentally tested. Thus, similar results are demonstrated for the comparison between lines and slots. Differences are observed concerning the increase of the width of the oxide hole. Nevertheless, the sizes of the holes considered are different: $3\mu\text{m}$ vs. $6\mu\text{m}$ for the study from Fiori et al. and $1\mu\text{m}$ vs. $2\mu\text{m}$ for the current study, and intermediate sizes of holes are maybe critical for the pad reliability.

In their studies, Rhee et al. [Rhee, 2003] and Vayrette et al. [Vayrette, 2010] demonstrated through experiments and modelling that the copper in small lines, i.e. in the range from $0.25\mu\text{m}$ to $2\mu\text{m}$, have non neglecting residual stress due to the thermal stages of the die process. The residual stress is found to be tensile stress, in the range of few hundreds of mega Pascal. Moreover, they found that increasing the line width is related to a decrease of the residual stress. For $2\mu\text{m}$ lines, the latter underlined that the stress is close to the one of a thin film. Additionally, Muppidi and co-workers [Muppidi, 2005] evaluate by means of EBSD observations and analyses, the effect of the line width on the copper grain size filling it. They showed that increasing the line width is related to a small increase of the average grain size (between $1\mu\text{m}$ and $2\mu\text{m}$). A variation of the grain orientation is observed as well. Then, very

high differences in material properties (and the Young's modulus particularly) are found for different copper grain orientations. The yield stress is also found to increase while decreasing the line width. In spite of the field of study restricted to copper lines, similar behaviour can be expected with slotted copper plates, which can explain the differences observed during the experiments. It can also explain why different results are found in the literature [Fiori, 2007] since constant elastic isotropic material properties of the copper are used in the numerical model.

IV. Nanoindentation and modelling to reproduce EWS and wire bonding

In this section, the correlations between the techniques used to try to mimic the probing and wire bonding are discussed. On the one hand, the finite element model is set up to simulate the EWS test. Modelling allows evaluating a large amount of designs at reduced cost, since the patterned wafers are not required to compare the structures. The Von Mises and compressive stresses are evaluated in all the layers and a particular attention is paid to the oxide ones, within the cracks are experimentally observed. For the thicknesses effects, the same trends are found as the actual tests, including the cumulative effect while increasing the thicknesses of aluminium and/or pad open oxide. Nevertheless, due to the use of the homogenization technique, the layout of the top copper layer provides the good results only when different structures are investigated, i.e. different orthotropic material properties are required to weight the pads. Among other, submodeling can be used to try to discriminate the structures.

On the other hand, the nanoindentation measurements with Berkovich and cube corner tips provide complementary results. The nanoindentation is useful to evaluate the pads much easily than the probing, by applying a compressive loading. Even though wafers are needed to perform the tests, neither the prober, nor the expensive probe card is required. Indenting with the Berkovich tip has reproduced the same failures at geometrical singularities, and the same plastic behaviours of the aluminium and copper layers as wafer sorting. Both tips give complementary results, and the cube corner tip allows a more local evaluation of the materials. Initially, the depth reached at maximum load is used to compare the structures. Moreover, the cube corner tip is able to generate cracks in the pad open oxide. These failures correspond to an energy release characterized by a discontinuity in the load-displacement curves. As a consequence, a new indicator ($h \cdot dh/dp$) is plotted and used to obtain the critical load and displacement into the surface prior to failure. Then, all these results are combined to rank the structures according to their mechanical robustness. The effect of the thicknesses of the aluminium and oxide layer is reproduced. Nevertheless, due to the ease of the tips to penetrate through the aluminium layer, the nanoindentation provide good results for the effects separated, i.e. one parameter modified at a time, and cross comparison do not depict the right results. Still on the same CMOS 120nm samples, the evaluation of the top copper design is able to reproduce most of the tests results. Only the shift of the plate has not been obtained. This design effect is a global effect that requires taking into account the shear stress and evaluating global response of the pad. Finally, for the study of the EWS nanoindentation and the numerical model have shown their ability to reproduce the results, even if few limitations have been found in each case.

Then, the nanoindentation measurements are compared to the wire pull tests. The loading condition in the two cases is quite different, since the wire bonding induces not only compressive loading, but also important shear and tensile loads during the process or the pull test. For these particular samples, the nanoindentation has not been able to predict the peeling results. It is related to the presence of the geometrical singularities in the pad open via design, where the peeling occurs. Their presence enhances the resistance to compression, but reduces

the resistance to shear. However, for some other pads without such singular points, nanoindentation is able to obtain the same results as the wire pull tests¹. Then, for the pads study described in this chapter, it seems that a failure threshold is obtained, explaining the failures and following peeling cases. Nanoindentation is also able to predict the resistance to compression for latest technology nodes with low-k oxide, and can be used to strengthen the mechanical properties of these layers.

Finally, both the finite element modelling and the nanoindentation are shown to be able to reproduce the results of the tests and processes performed on the pad during the manufacturing steps. These tools are very efficient to forecast further failures in the pads and can be used for technology development or for crisis resolution, considering the few limitations aforementioned.

V. Conclusions

In this chapter, most of the currently available designs to strengthen the pad robustness have been probed or simulated. In a first step, many EWS tests, wire pull tests and related characterizations have been performed to gain insight on the failure mechanisms, and to obtain baseline for further comparison. Hence, the behaviour of aluminium and copper during these tests and their respective role on the robustness of the pad is underlined. The aluminium acts as a stress buffer layer to avoid the damages in the interconnections. These damages are cracks or peeling, which occur in the pad open oxide or inter-metal dielectric layers. The geometrical singularities in the pad have been pointed out to be detrimental to the pad robustness while submitted to such tests. Then, the finite element model and the nanoindentation have been used to mimic these test and process. The development of indicators for each analysis technique has been achieved to provide reliable results for the pad study.

On the pad designs effects, all the techniques show that increasing the thicknesses of the aluminium and pad open oxide layers is related to a reduction of the failure occurrence in the pad open oxide and inter-metal dielectric layers. Then, the metal density of the copper layer is demonstrated to have an effect on the mechanical robustness of the pad. However, the increase or the decrease of the copper density is intimately linked to the oxide materials surrounding it. Indeed, the differences in materials properties between the copper and either the silicon glass or the low-k oxide is consequent. Thus, for technology nodes using the former oxide, the reduction of the metal density increases the pad resistance to compressive loading. For latest nodes, using the low-k oxide, the opposite trend is advised to strengthen the pad. The presence of via between the copper layers has not been related to any variation of the mechanical robustness of the pads. Neither nanoindentation nor wire pull and shear tests have underlined this trend. The presence of the pad open oxide layer is observed to improve the resistance to compression, however, while via are inserted within the layer, and under the probing or bonding area, the numerous geometrical singularities introduced as well are shown to increase dramatically the number of peeling cases. The layout of the top copper level can be modified in order to reduce the fracture hazard. As a matter of fact, the pad layout with slotted plates, and particularly with smallest oxide holes, should be preferred to wide copper lines while designing the pad. The location of the plates with respect to the probing location is also able to reduce the crack occurrence.

Finally most of the pad layout parameters have an effect on the mechanical robustness of the pad, and their modification according to the results previously described can help to

¹ See appendix A for additional details one these experiments

reduce the mechanical issues in the pad. Additionally, a methodology to evaluate the pads has been given in this chapter. The use of nanoindentation and finite element modelling is found to be able to reproduce most of the probing and some of the wire bonding issues. Hence, these tools can be used for the technology development and to solve crises by giving fast and reliable alternatives to the costly actual tests and processes operations applied to the pad. Yet, for the nanoindentation, choices of the tip for a given layout and pad study methodologies have to be clarified to perform reliable tests. Similarly, a less CPU consuming finite element model to mimic EWS is required in the industrial context to evaluate more rapidly the various pad options.

References

- [Beegan, 2005]** Beegan, D., Chowdhury, S. and Laugier, M.T., Work of indentation methods for determining copper film hardness, *Surface and coatings technology*, vol. 192, 2005, pp. 57-63.
- [Brillet-Rouxel, 2006]** Brillet-Rouxel, H., Arfan, E., Leguillon, D., Dupeux, M., Braccini, M. and Orain, S., Crack initiation in Cu-interconnect structures, *Microelectronics engineering*, vol. 83, 2006, pp. 2297-2302.
- [Cecchetto, 2007]** Cecchetto, L., Zullino, L., Vallauri, R., Andreini, A., Redaelli, L., Anghilieri, P. and Vettori, R., Mechanical simulation of electrical wafer sort using pad over active structures in 0.18 μ m smart power technology, Excerpt from the proceedings of the COMSOL users conference, 2007.
- [Chen, 2005]** Chen, S., Liu, L. and Wang, T., Investigation of the mechanical properties of thin films by nanoindentation, considering the effects of thickness and different coating substrate combinations, *Surface and coatings technology*, vol. 191, 2005, pp. 25-32.
- [Chen, 2006]** Chen, J. and Bull, S.J., Assessment of the toughness of thin coatings using nanoindentation under displacement control, *Thin solid films*, vol. 494, 2006, pp. 1-7.
- [Chin, 2007]** Chin, P.-C., Hu, C.H., Hsu, H.-C., Fu, S.-L., Yeh, C.-L. and Lai, Y.-S., Characteristic of heat affected zone in thin gold wire and dynamic transient analysis of wire bonding for microstructure of cu/low-k wafer, *International microsystems, packaging, assembly and circuit technology, IMPACT 2007*, pp. 297-300.
- [Degryse, 2004]** Degryse, D., Vandeveld, B. and Beyne, E., Mechanical FEM simulation of bonding process on cu lowk wafers, *IEEE transactions on components and packaging technologies*, vol. 27, no. 4, December 2004, pp. 643-650.
- [Doerner, 1986]** Doerner, M.F. and Nix, W.D., A method for interpreting the data from depth-sensing indentation instruments, *Journal of material research* vol. 1, issue 4, July/August 1986, pp. 601-609.
- [Engelen, 2007]** Engelen, R.A.B., van des Sluis, O., van Silfhout, R.B.R., van Driel, W.D. and Fiori, V., Optimization of cu low-k bond pas designs to improve mechanical robustness using area release energy method, *IEEE international conference on thermal, mechanical and multiphysics simulation and experiments in microelectronics and microsystems, Eurosime 2005*, pp. 1-4.
- [Fiori, 2005]** Fiori, V., Orain, S., A multi-scale finite element methodology to evacuate wire bond pad architectures, *IEEE international conference on thermal, mechanical and multiphysics simulation and experiments in micro-electronics and micro-systems*, 2005, pp. 648-655.
- [Fiori, 2007]** Fiori, V., Beng, L.T., Downey, S., Gallois-Garreignot, S. and Orain, S., 3D multi scale modeling of wire bonding induced peeling in cu/low-k interconnects: application of an energy based criteria and correlations with experiments, *IEEE electronics components and technology conference*, 2007, pp. 256-263.
- [Fischer-Cripps, 2006]** Fischer-Cripps, A.C., Critical review of analysis and interpretation of nanoindentation test data, *Surface and coatings technology*, vol. 200, 2006, pp. 4153-4165.

- [Fujikane, 2008]** Fujikane, M., Nagao, S. Liu, X.W., Chrobak, D., Lehto, A., Yamanaka, S. and Nowak, R., Evaluation of carbon-doped low-k multilayer structure by nanoindentation, *Journal of alloys and compounds*, vol. 448, issues 1-2, 2008, pp. 293-297.
- [Gallois-Garreignot, 2007]** Gallois-Garreignot, S., Fiori, V., Orain, S. and van der Sluis, O., Numerical analysis of the reliability of cu/low-k bond pad interconnections under wire pull test: application of a 3D energy based failure criterion, *IEEE international conference on thermal, mechanical and multiphysics simulation and experiments in micro-electronics and micro-systems*, 2007, pp. 1-7.
- [Hay, 2000]** Hay, J.L. and Pharr, G.M., *Instrumented indentation testing*, Material Park, ASM International, 2000, pp. 232-243.
- [Hay, 2009]** Hay, J., Measuring substrate-independent modulus of dielectric films by instrumented indentation, *Journal of materials research*, vol. 24, no. 3, March 2009, pp. 667-677.
- [Hsu, 2007]** Hsu, H.C., Chang, W.Y, Fu, S.L., Yeh, C.L. and Lai, Y.S., Dynamic finite element analysis on underlay microstructure of cu/low-k wafer during bonding process, *International conference on electronic materials and packaging*, 2007, pp. 1-6.
- [Jang, 2008]** Jang, J.-I. and Pharr, G.M., Influence of indenter angle on cracking in Si and Ge during nanoindentation, *Acta Materialia*, vol. 56, 2008, pp. 4458-4469.
- [Kregting, 2006]** Kregting, R., van Silfhout, R.B.R., van der Sluis, O., Engelen, R.A.B., van Driel, W.D. and Zhang, G.Q., An experimental-numerical study of metal peel off in cu/low-k back-end structures, *International conference on electronics packaging technology*, 2006, pp. 1-6.
- [Lee, 2007]** Lee, C.C., Tran; T.A. and Miller, C., Overview of metal lifted failure modes during fine pitch wirebonding low-k/copper dies with bond over active circuitry design, *IEEE electronics components and technology conference*, 2007, pp. 1775-1781.
- [Liu, 2005]** Liu, Y, Desbiens, D., Irving, S. and Luk, T., Probe test failure analysis of bond pad over active structure by modeling and experiment, *IEEE electronic components and technology conference*, 2005, pp. 861-866.
- [Liu, 2007]** Liu, Y., Liu, Y., Irving, S. and Luk, T., Wafer probing simulation for copper bond pad based BPOA structure, *IEEE international conference on thermal, mechanical and multiphysics simulation and experiments in microelectronics and microsystems*, Eurosime 2007, pp. 1-5.
- [Liu, 2008]** Liu, Y., Irving, S. and Luk, T., Thermosonic wire bonding process simulation and bond pad over active stress analysis, *IEEE transactions on electronics packaging manufacturing*, vol. 31, no. 1, January 2008, pp. 61-71.
- [Liu, 2008]** Liu, Y., Liu, Y., Irving, S., Luk, T., Wang, Q., Investigation of BPSG profile and FAB size on cu stub bumping process by modeling and experiments. *IEEE international conference on thermal, mechanical and multiphysics simulation and experiments in micro-electronics and micro-systems*, 2008, pp. 312-317
- [Mercado, 2000]** Mercado, L., Radke, R., Rutson, M., Tran, T.A., Williams, B., Yong, L., Chen, A., and Chen, S., Reliability of multi-layer aluminium capped copper interconnect structures, *IEEE International electronics manufacturing symposium*, 2000, p. 84-93.
- [Morris, 2008]** Morris, D.J. and Cook, R.F., Indentation fracture of low-k dielectric constant films: part I. Experiments and observations, *Journal of materials research*, vol. 23, 2008, pp. 2429-2442.

- [Muppidi, 2005]** Muppidi, T., Field, D.P., Sacher, J.E. and Woo, C., Barrier layer, geometry and alloying effects on the microstructure and texture of electroplated copper thin films and damascene lines, *Thin solid films*, vol. 471, 2005, pp. 63-70.
- [Nay, 2004]** Nay, R.J., Warren, O.L., Yang, D. and Wyrobek, T.J., Mechanical characterization of low-k dielectric materials using nanoindentation, *Microelectronics engineering*, vol. 75, 2004, pp. 103-110.
- [Ng, 2006]** Ng, B.T., Ganesh, V.P. and Lee, C., Optimization of gold bonding on electroless nickel immersion gold for high temperature applications, *IEEE electronic packaging technology conference*, 2006, pp. 277-282.
- [Oliver, 1992]** Oliver, W.C. and Pharr, G.M., An improved technique for determining hardness and elastic modulus using load and displacement sensing indentation experiments, *Journal of materials research* 1992, vol. 7 (6), p. 1564-1583.
- [Rhee, 2003]** Rhee, S.-H., Du, Y. and Ho, P.S., Thermal stress characteristics of cu/oxide and cu/low-k submicron interconnect structures, *Journal of applied physics*, vol. 93, number 7, 2003, pp.3926-3933.
- [Saha, 2002]** Saha, R. and Nix, W.D., Effects of the substrate on the determination of thin film mechanical properties by nanoindentation, *Acta materialia*, vol. 50, 2002, pp. 23-38.
- [Simonyi, 2005]** Simonyi, E.E., Liniger, E., Lane, M., Lin, Q., Dimitrakopoulos, C.D. and Tyberg, C., Application of nanoindentation to characterize fracture in ILD films used in the BEOL, *Proceeding of the materials research society symposium*, vol. 863, 2005, pp. 139-144.
- [Soare, 2004]** Soare, S., Bull, S.J., O'Neil, A.G., Wright, N., Horsfall, A. and dos Santos, J.M.M., Nanoindentation assessment of aluminium metallisation; the effect of creep and pile-up, *Surface and coatings technology*, vol. 177-178, 2004, pp. 497-506.
- [Tran, 2000]** Tran, T.A., Yong, L. Williams, B., Chen, S. and Chen, A., Fine pitch probing and wirebonding and reliability of aluminium capped copper bond pads, *IEEE electronics components and technology conference*, 2000, pp. 1674-1680.
- [van der Sluis, 2006]** van der Sluis, O., Engelen, R.A.B, van Silfhout, R.B.R, van Driel, W.D. and van Gils, M.A.J., Efficient damage sensitivity analysis of advanced cu low-k bond pad structures using the area release energy criterion, *IEEE international conference on thermal, mechanical and multiphysics simulation and experiments in micro-electronics and micro-systems*, 2006, pp. 1-8.
- [van Gils, 2005]** van Gils, M.A.J., van der Sluis, O., Zhang, G.Q., Janssen, J.H.J. and Voncken, R.M.J., Analysis of cu-low-k bond pad delamination by using a novel failure index, *IEEE international conference on thermal, mechanical and multiphysics simulation and experiments in micro-electronics and micro-systems*, 2005, pp. 1-7.
- [Vayrette, 2010]** Vayrette, R., Rivero, C., Gros, B., Blayac, S. and Inal, K., Residual stress estimation in damascene interconnects, using embedded sensors, *Microelectronics engineering*, vol. 87, pp. 412-415.
- [Vella, 2003]** Vella, J.B., Adhihetty, I.S., Junkker, K. and Volinski, A.A., Mechanical properties and fracture toughness of organo-silicate glass (OSG) low-k dielectric thin films for microelectronic applications, *International journal of fracture*, vol. 119-120, 2003, pp. 487-499.
- [Viswanath, 2007]** Viswanath, A.G.K., Zhang, X., Ganesh, V.P. and Chun, L., Numerical study of gold wire bonding process on cu/low-k structures, *IEEE transactions on advanced packaging*, vol. 30, no. 3, August 2007, pp. 448-456.

[Yeh, 2006] Yeh, C.L. and Lai, Y.S., Comprehensive dynamic analysis of wirebonding on cu/low-k wafers, IEEE transactions on advanced packaging, vol. 29, no. 2, May, 2006, pp. 264-270.

[Wu, 2006] Wu, F., Gonchond, J.-P., Optimisation de la technique de nanoindentation pour la caractérisation des propriétés mécaniques des oxydes basse permittivité, Université Joseph Fourier, Grenoble, France, Master degree thesis, 2006.

Chapter 5:

Enhancement of the numerical and experimental methods and tools to evaluate the robustness of pads

Outline

I.	Nanoindentation: qualification for pad robustness study	- 164 -
I.A.	Comparison between Berkovich and cube corner tips	- 164 -
I.A.1.	Dimension and shape of the indent	- 164 -
I.A.2.	Influence of the tip geometry on the load-displacement curves.....	- 165 -
I.A.3.	Determination of the Young's modulus and hardness	- 166 -
I.B.	Application to the determination of the pad robustness	- 168 -
II.	Modelling: FEM dedicated to pad enhancement to reproduce EWS	- 169 -
II.A.	Comparison between explicit and implicit methods.....	- 169 -
II.A.1.	Time integration	- 169 -
II.A.2.	Stability limits.....	- 170 -
II.B.	Development of the 2D implicit model	- 170 -
II.B.1.	Objectives and use of experimental and numerical insights.....	- 170 -
II.B.2.	Model description.....	- 171 -
II.B.3.	Loading conditions of the 2D model	- 172 -
II.B.4.	Results and limitations of the 2D FE model.....	- 176 -
II.C.	Development of the 3D implicit FE model.....	- 179 -
II.C.1.	Description of the 3D FE model of the pad.....	- 179 -
II.C.2.	Loading conditions	- 179 -
II.C.3.	Results of the 3D model	- 180 -
II.C.4.	Insertion of an elastic-plastic behaviour for the aluminium.....	- 182 -
II.D.	EWS simulations: explicit versus implicit 3D modelling.....	- 187 -
III.	Conclusions	- 187 -
	References	- 189 -

In the previous chapters, the mechanical issues in the pads during EWS and the wire bonding followed by wire pull and shear tests have been investigated. In parallel, nanoindentation and finite element modelling have been used to evaluate the mechanical robustness of the structures and rank them. Then, some correlations between the various techniques have been demonstrated, and nanoindentation and the probing modelling have been able to reproduce the failures and trends from the actual tests. However, in the industrial context, fast and reliable analysis techniques are required. More precisely, for the nanoindentation technique, guidelines for the pad study are necessary to compare their mechanical robustness. In addition, a less CPU time consuming and pad's enhancement dedicated numerical model also need to be set up.

Indeed, in the chapter 4, distinct nanoindentation results have been obtained using the Berkovich and the cube corner tips. Hence, it appears necessary to clarify test guidelines, such as the loading condition to apply to the pad. Some indications on the tip selection are also required since the choice depends on the pad structure to evaluate.

Then, a 3D finite element model needs to be developed. It must take into account all the experimental and numerical observations and results already obtained to define accurate loading conditions mimicking the probing. In the frame of reducing the CPU time consumption, the simulation of the contact between the probe and the pad should be avoided. Thus, specific loading conditions, based on the Hertz contact, are applied and their ability to reproduce the actual EWS loading conditions is verified by introducing elastic plastic material properties for the aluminium. Afterwards, the effects of the overdrive and the number of passes on the scrub size are evaluated.

Finally, a full methodology, including both experimental and numerical methods, has been investigated for pad yield optimization. Hence, an efficient tool set for manufacturing process developments is proposed.

I. Nanoindentation: pad robustness qualification

In the previous chapter, the Berkovich and the cube corner tips have been used to evaluate the pads. Usually, most of the trends observed are similar for the two tips. Hence, in this part, the differences between the indenters are discussed, and then, general rules for the evaluation of the pad's mechanical robustness are provided.

I.A. Comparison between Berkovich and cube corner tips

I.A.1. Dimension and shape of the indent

The geometrical considerations of the Berkovich and the cube corner tips have been described in the section II.A.2 of the chapter 2. Thus, the projected area and the volume-depth relation are lower for the cube corner tip. As a consequence it has been underlined in the previous chapter that for an equivalent depth, the indent left in the pad is much bigger with a Berkovich tip. Yet, cracks are more easily induced in the oxide with the cube corner tip.

The Figure V.1 depicts the projected area of the two tips. The pile-up is larger in case of indentation with the Berkovich tip. However, when related to the indent size, the cube corner tip induces a more important pile-up. For these indents, the pile-up heights measured during the SEM observations after FIB milling are respectively in the order of 0.6µm and 1µm for the cube corner and Berkovich tips.

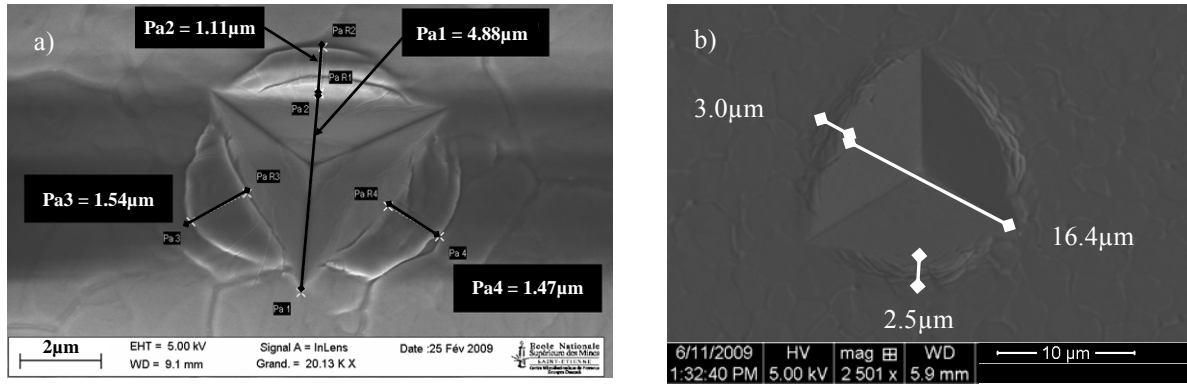


Figure V.1. SEM top views of the pads after nanoindentation with cube corner (a) and Berkovich (b) tips on aluminium.

I.A.2. Influence of the tip geometry on the load-displacement curves

The load-displacement curves for a given penetration are shown in the Figure V.2. Thus, much lower load is required to reach the same depth while using the cube corner tip. In the literature, many works have highlighted similar differences for the load-displacement curves obtained from the two indenter tips. [Bucaille, 2004-A; Bucaille, 2004-B; Mirshams, 2004; Johnson, 2007]. More generally, for a three-sided punch, the load required to reach a given depth decreases while decreasing the equivalent cone angle [Chollacoop, 2003; Jang, 2008].

The load-displacement curves described are normalized (over their maximum load and displacement) to compare the penetration of the two tips in the multi-layers of the pad. As depicted in the Figure V.3, the Berkovich tip has more a parabolic evolution of the displacement while increasing the load. On the contrary, the cube corner tip induces a clear slope break between the aluminium and the oxide layers. This slope break occurs almost at the interface between the two materials, which underlines the local sollicitation of the material parameters with the latter indenter. The unloading curve obtained from indentation with the two tips is also different. Indeed, higher elastic recovery is observed for the Berkovich tip, i.e. the h_f/h_{max} ratio is lower.

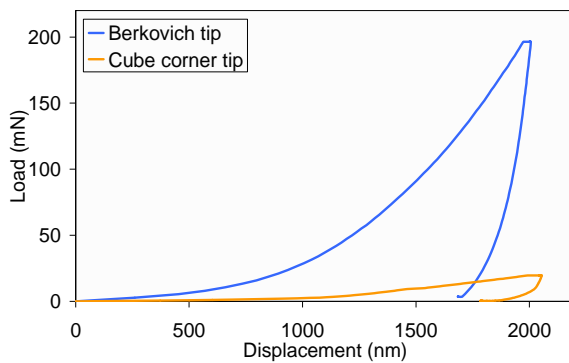


Figure V.2. Load-displacement curves for typical indentations with Berkovich and cube corner tips on CMOS 45nm samples.

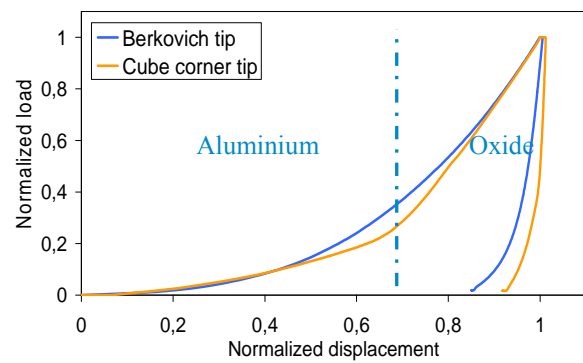


Figure V.3. Normalized load versus normalized displacement curves for indentations with Berkovich and cube corner tips.

The Table V.1 summarizes the results obtained during the study on the CMOS 45nm samples. These pads have been studied in the section II.C of the chapter 4. The pad layout 6 is evaluated (metal density: 70%, with via and CB via). Higher standard deviation is observed for the cube corner tip. It implies a reduction of the measurement stability, particularly while

the results between the structures are close. The higher variation between the measurements is due to the crack observed on the load-displacement curve for the cube corner tip. Additionally, the local stresses and strains induced by this tip imply a higher dependence on the small details (e.g. pad open vias) met during the tip penetration.

	Berkovich tip	Cube corner tip
Average maximum displacement (in nm)	2194	2162
Standard deviation (in nm)	18	64
Uncertainty (std deviation/max. displacement)	0.8%	3.0%
Elastic recovery (h_f-h_{max} , in nm)	318	162
h_f/h_{max}	0.86	0.93

Table V.1. Nanoindentation results for both indenter tips on CMOS 45nm samples.

In the literature, the higher stresses and strains induced by the cube corner tip have been evaluated by Bucaille et al. [Bucaille, 2003]. They have shown from indentation simulations on copper that for a small included angle (i.e. a cube corner tip), the plastic strains are localized near the indenter. Moreover, the plastic volume deformed is higher for the Berkovich indenter than for the cube corner one. Nevertheless, plastic strains are higher in the case of indentation with cube corner tip, as depicted in Figure V.4.

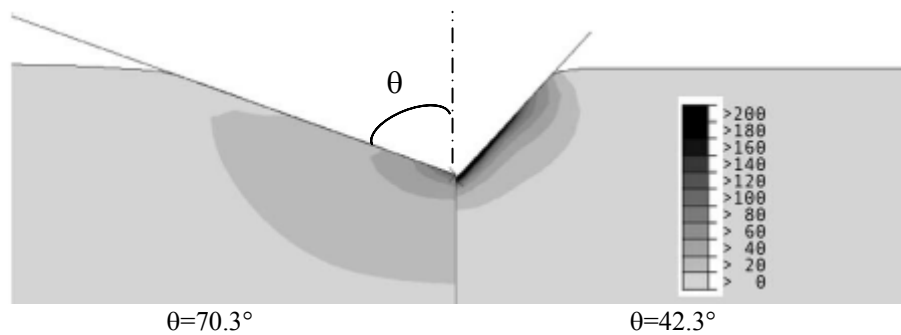


Figure V.4. Map of equivalent plastic strains in copper near the tip for Berkovich and cube corner tips, after [Bucaille, 2003].

Similarly, through finite element simulations, Durst and co-workers [Durst, 2006] evaluated the plastic zone left by an indentation with a cube corner tip. They described the extension of the plastic zone, and compared it for indentations with the Berkovich and the cube corner tips.

I.A.3. Determination of the Young's modulus and hardness

In order to have reliable inputs in the numerical models, some nanoindentation with the Berkovich and the cube corner tips have been made on a copper film, whose thickness is $1.95\mu\text{m}$, deposited on a silicon substrate. Hence, some measurements with the CSM mode are performed to obtain the material properties as a function of the displacement into the pad.

The depth reached at the end of the loading step is 1400nm , which is lower than the film thickness. The loads applied to reach this depth are 78.7mN and 7.96mN for the Berkovich and the cube corner tips respectively.

The normalized load-displacement curves with the two indenters are shown in the Figure V.5. The parabolic shape described in several papers [Chollacoop, 2003; Bucaille 2003] for

the indentation with the Berkovich tip is observed. Indeed, the following relationship is obtained:

$$P = C * h^2 \quad (V.1)$$

where the value of C is 3.79 for the current indentation in copper.

During the unloading stage, the elastic recovery is more important for the indentation with the Berkovich tip. This trend on copper is similar to the one observed for the multi-layer stacks. It underlines the effect of the cube corner tip geometry, which induces higher stresses and strains to the copper.

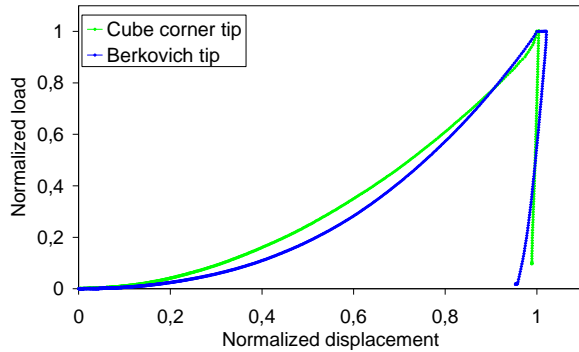


Figure V.5. Normalized load-displacement curves for nanoindentation in a thick copper film on silicon.

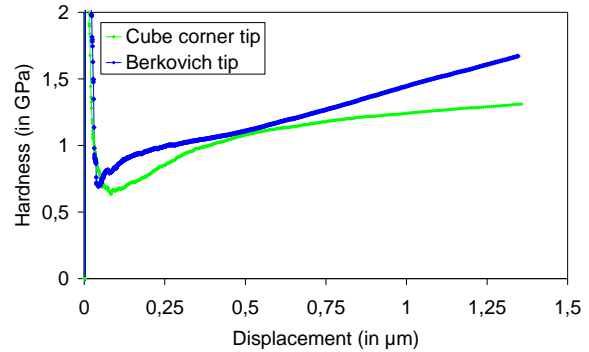


Figure V.6. Hardness as a function of the displacement for the Berkovich and the cube corner tips.

Then, the Young's modulus and the hardness of the copper are determined. The material properties obtained are respectively 132.9GPa and 1.1GPa for the Berkovich tip and 127.7GPa and 1GPa for the cube corner one. The Young's modulus and the hardness deduced from the Berkovich tip are slightly above these of the cube corner one. As aforementioned, this difference is related to the presence of the important pile-up for the latter tip. Then, the Figure V.6 shows that the hardness measured with the Berkovich tip is, whatever the displacement, slightly above the one deduced from the cube corner tip. Moreover, while increasing the indentation depth, the difference between the two tips increases. It is explained by the presence of the silicon below the copper layer, whose Young's modulus (166GPa) and hardness (about 8GPa) are above these of the copper. The high hardness values observed at the shallow indentations depth are referred in the literature as the indentation size effect (ISE), which is explained by the presence of geometrically necessary dislocations [Nix, 1998]. Then, the low values at the beginning of the curves are probably explained by the presence of a thin (40nm) Si₃N₄ barrier layer on top of the copper. This material has higher hardness than copper, but in this stack, it is a very thin and fragile layer.

These measurements are in agreement with these made in the literature. Actually, hardness measurements from the Berkovich tip are found below the ones from the other tip for measurements on micro and nano-crystalline pure nickel [Mirshams, 2004; Mirshams, 2006], on poly-crystalline nickel and copper [Durst, 2006]. Yet, in other works, the hardness measurements on polymethyl methacrylate, stainless steel, copper, iron and aluminium [Wei, 2008] or in low-k oxides (BD2X) [Wu, 2006] are of the same range.

Finally, small measurement variations of the mechanical properties are also obtained while comparing the Berkovich and the cube corner tips. It underlines the difference between the geometries of the two indenters.

I.B. Application to the determination of the pad robustness

In the chapter 4 and in the previous sections, some differences have been shown while comparing the Berkovich and the cube corner tips. Hence, some standard rules for the analysis of pads, depending on its design and the expectations are given.

The nanoindentation maximum displacement into the surface is expected to reach about 2 layers to obtain quite an overview of the pad's robustness. While oxide is detected below the top layer, nanoindentation has to penetrate most of it in order to try to generate a crack within the layer. Much deeper nanoindentation into the pad may induce additional interpreting issues due to the multi-layer systems and should be avoided. Thus, for the CMOS 45nm and the CMOS 120nm pad, a depth of 2 μ m is good to obtain a ranking between the structures. For the technology nodes in between, the same maximum displacement into surface is expected to provide good results as well.

Then, several tests are required for reproducibility purposes. Actually, between 15 and 30 indentation tests on a given structure strengthen the results reliability. Generally, fewer tests are required with the Berkovich tip due to its higher reproducibility.

The two tips have to be used if possible, since they provide complementary information on the pad's mechanical robustness. Indeed, during the study on pads, the Berkovich tip is able to reproduce the cracks induced by the EWS test, and in some cases the loading conditions undergone by the pad during the wire bonding process. Yet, the cube corner tip induces high stresses and strains in the materials, which lead to the appearance of noticeable cracks in the pad open oxide layers. Whenever the use of the two tips is not possible, the flow chart depicted in the Figure V.7 provides help to choose the tip that best suit with the pad study (and architecture). The tip selection can be necessary due to lack of time or lack of samples to perform the analysis.

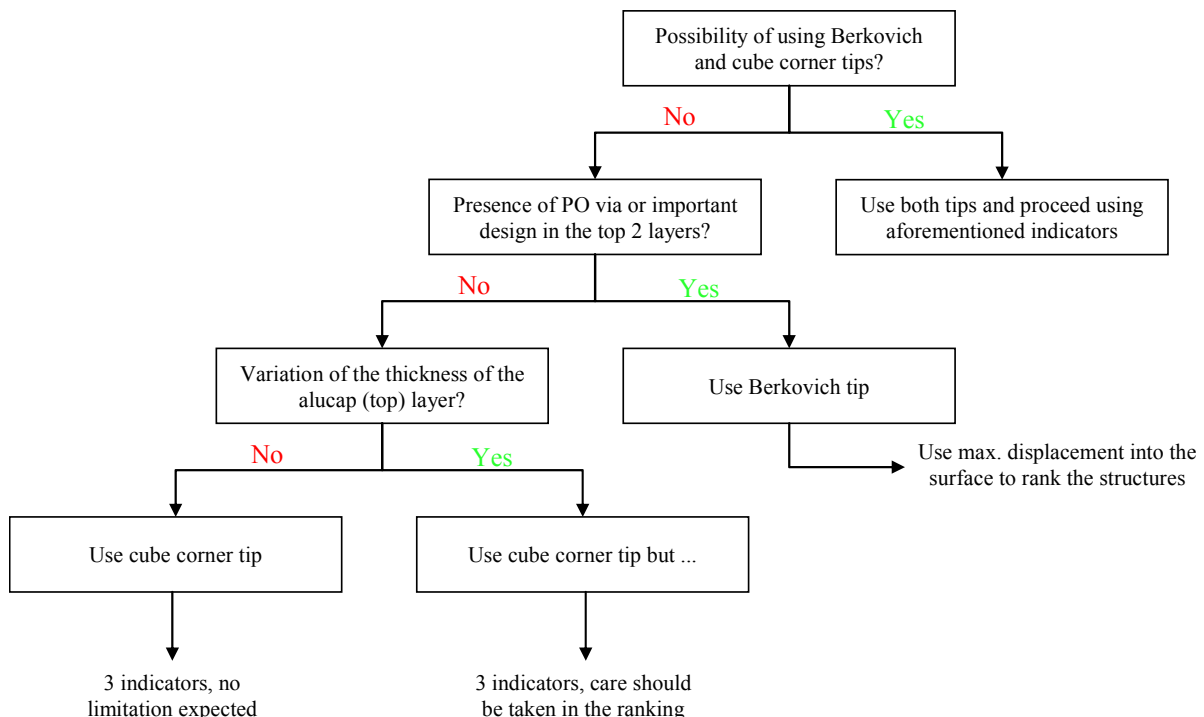


Figure V.7. Flow chart of the nanoindentation on pads: tip selection as a function of the investigated layout.

II. Modelling: FEM dedicated to pad enhancement to reproduce EWS

In the previous chapter, the explicit numerical model is developed to mimic the probing. For most of the test and design parameters, the numerical model has been able to predict the EWS results, i.e. rank the structures according to their mechanical robustness. However, the model is highly CPU time consuming since about one hundred hours are required to obtain the results from one simulation (i.e. one stack). Moreover, some limitations have been highlighted for the pad study due to the homogenization technique. In the industrial context, and more specifically, in the development of new technologies or during the crises resolution, several pad structures need to be simulated and thus, a very long time is required to obtain the results. Hence, in this part, a new finite element model has to be developed to compare the structures more rapidly. This model has to take into account all the insights collected from the various experiments and simulation previously made.

This model consists only in the pad. Thus, the contact between the probe tip and the pad is not represented, and then the numerical related non-linearity and convergence issues (e.g. the CMOS 45nm calibration) do not need to be modelled.

II.A. Comparison between explicit and implicit methods

The finite element models of the EWS have been solved using an explicit scheme (chapters 2, 3 and 4). In this chapter, an implicit scheme is considered. In the following section, these methods and the related assumptions and applications are described. Deeper comparison of the two methods is given in the Ansys user manual [Ansys, rel. 11].

II.A.1. Time integration

II.A.1.a) Implicit time integration

In the implicit time integration method, the inertia effects of mass and damping ($[C]$ and $[M]$) are not included. The average acceleration and displacements evaluated at time $t+\Delta t$ are given by:

$$\{u_{t+\Delta t}\} = [K]^{-1} \{F_{t+\Delta t}^a\}$$

Hence, for linear problems, the solution is unconditionally stable when the stiffness matrix $[K]$ is linear and large time step can be chosen. For non linear problems, the solution, which requires the inversion of the non linear stiffness matrix, is obtained using a series of linear approximations (Newton-Raphson method). However, small iterative steps are necessary to achieve the convergence, which is not always obtained, especially for highly nonlinear problems.

II.A.1.b) Explicit time integration

In the explicit time integration method, the accelerations are evaluated at a time t , following:

$$\{a_t\} = [M]^{-1} (\{F_t^{ext}\} - \{F_t^{int}\})$$

where $\{F_t^{ext}\}$ is the applied external and body force vector and $\{F_t^{int}\}$ is the internal force vector, which is given by:

$$F^{int} = \sum \left(\int_{\Omega} B^T \sigma_n d\Omega + F^{hg} + F^{con} \right)$$

where F^{hg} is the hourglass resistance force and F^{con} is the contact force.

Then, the velocity and displacements are evaluated:

$$\begin{aligned}\{v_{t+\Delta t/2}\} &= \{v_{t-\Delta t/2}\} + \{a_t\}\Delta t_t \\ \{u_{t+\Delta t}\} &= \{u_t\} + \{v_{t+\Delta t/2}\}\Delta t_{t+\Delta t/2}\end{aligned}$$

where $\Delta t_{t\pm\Delta t/2} = 0.5(\Delta t_t \pm \Delta t_{t+\Delta t})$

The geometry is updated by adding the displacements increments to the initial geometry $\{x_0\}$:

$$\{x_{t+\Delta t}\} = \{x_t\} + \{u_{t+\Delta t}\}$$

As a consequence, the equations become uncoupled and can be solved directly, even for nonlinear problems. Since no inversion of the stiffness matrix is required, all the nonlinearities and the contact are included in the internal force vector. Finally, no convergence issues are faced because the equations are uncoupled. Nevertheless, very small time steps are required to maintain stability limit, which implies high computation duration.

II.A.2. Stability limits

II.A.2.a) Implicit time integration

As previously mentioned, the implicit solution is always stable and the time step can be arbitrarily large. For nonlinear problems, the time step size may become small due to convergence difficulties.

II.A.2.b) Explicit time integration

The explicit solution is only stable if the time step is smaller than the critical time step size (as noticed in the section III.A.4.b) of the chapter 2):

$$\Delta t \leq \Delta t^{crit}$$

As a consequence of the very small step size, the explicit method is only useful for very short transient phenomena.

II.B. Development of the 2D implicit model

II.B.1. Objectives and use of experimental and numerical insights

In the chapters 2 to 4, observations, experimental and numerical studies have provided much information on the probing and its related failures in the pad. These insights are used in the implicit model to define accurate loading conditions. The two dimensional implicit FE model do not contain the contact between the probe tip and the pad. Though, the loading conditions must be reproduced as fairly as possible.

In order to reproduce the test parameters (overdrive and number of passes), the loading conditions of the finite element model must depend on these parameters. In that frame, the measurements of the vertical and lateral reaction forces performed during the calibration step are used. Moreover, depending on the technology node, the pads pitch changes, and so does the probe card, the probe geometry... As a consequence, the numerical model must also take into account these modifications. The loading area, which represents the scrub mark left on the pad after testing, changes too, depending on the test conditions. Indeed, the results in the section I.B.1 of the chapter 3 have underlined the increase of the scrub length while increasing the overdrive and the limited effect of the number of passes.

Then, the shape of the loading has to refer to the experiments and simulations. As a matter of fact, during the explicit simulations, the contact between the probe tip and the pad has been demonstrated to be almost punctual due to the deflexion of the needle during the test (refer to

Figure IV.23). Finally, the failures have been observed in the oxide layers, and the largest one in the PO oxide, then, this model must end with similar stress peak location.

II.B.2. Model description

Only the pad is represented in the 2D implicit finite element model. The explicit model has shown some limitations to compare stacks due to the use of the homogenization technique. Then, it should be possible to introduce very small details such as vias in this model. The use of an implicit solver scheme enables us to use much more elements than the explicit model to describe the pad and thus the design. The model is developed using the Ansys release 11 software, and some information can be found in the user's manual [Ansys, rel. 11].

In this step, the portion of the pad represented consists in a cross section of the pad in the centre of the probe mark. Each layer is represented in the pad. In the model set up, full sheet layers with elastic isotropic material properties are used, as shown in the Table V.2. The 2D implicit finite element model describing the pad structure is depicted in Figure V.8. In order to be able to compare with previous results, the architecture of the CMOS 120nm pad is modelled. A 10 μ m thick silicon layer is added in the bottom of the model to avoid sharp boundary condition in the bottom of the pad. The mesh in this layer does not need to be as small as in the other layers since stresses in this layer are not the major issues.

Material	Young's modulus (GPa)	Poisson ratio
Aluminium	69	0.34
Oxide (TEOS)	63	0.23
Copper	128	0.33
Silicon	166	0.23

Table V.2. Material properties used in the 2D finite element model.

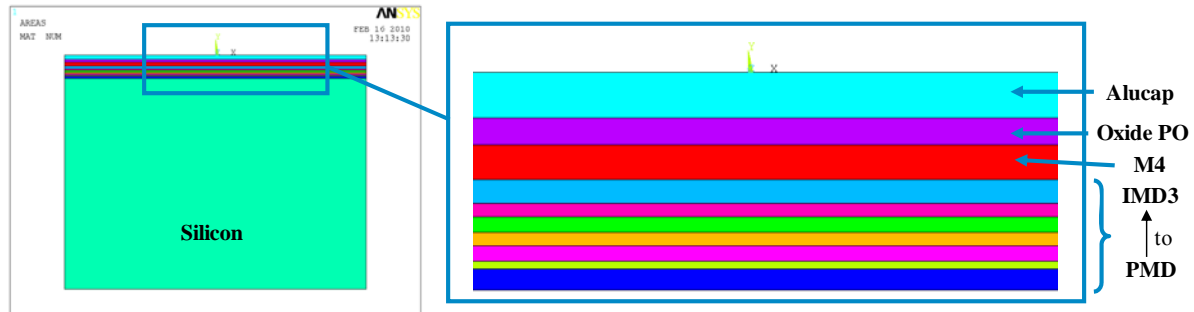


Figure V.8. 2D geometrical model of the CMOS 120nm pad.

The pad is meshed with some Plane183 elements, square shaped and quadratic, whose size is fixed at 0.2 μ m*0.2 μ m. By considering the characteristics of the pad, the plane strain assumption is made for these elements. This assumption is made while the third dimension is very important compared to the two represented in the model, and imposes the out of plane strains to be zero ($\epsilon_{zz} = \gamma_{xz} = \gamma_{yz} = 0$). Additional details and assumptions induced by the plane strain are described in a course by Lagace [Lagace, 2002]. In the literature, only a few papers [Ayhan, 1999; Krueger, 2002; Yeo, 2004] report studies between the 2D plane stress and plane strain options, compared to the 3D results, for the microelectronic field of applications. All the works showed that the results of the 2D plane stress and plane strain models are the boundaries of the true 3D model. At the package scale, they obtained results closer to the plane strain option. Besides, for the current model, since the mesh size is very small and no specific details are added to the pad, there is no meshing convergence issue.

Then, to mimic the probing loading, a pressure is applied to the pad. To that purpose, some SURF153 elements with mid-side nodes are added on its top. For these elements, it is possible

to apply normal and tangential pressures. These elements are used along a length that represents the length of the scrub mark left by the probe during EWS.

A set of boundary conditions is applied to the pad. On one lateral side, the vertical and the lateral Degrees Of Freedom (DOF) are removed for all the nodes ($u_x=0$ and $u_y=0$). On the other side of the pad, the displacement of all the nodes along the vertical is set to null ($u_y=0$) and the lateral displacement is constant ($u_x=\text{constant}$). At the bottom of the pad, both the vertical and lateral displacements are fixed at zero.

II.B.3. Loading conditions of the 2D model

Since the explicit model has underlined an almost punctual contact during the test and particularly at the end of the loading stage, the Hertz contact is introduced to reproduce the loading conditions of the probing. Then, to take into account the variations due to the technology nodes, the probe card, the needle geometry, etc, these parameters need to be in the modelling inputs. The analytical equations from the various calibration vertical and lateral reaction forces measurements depending on these parameters (and also obviously on the overdrive and the number of passes) are described to be finally applied to the pad.

II.B.3.a) Brief introduction to the Hertz contact

The contact is studied at the macroscopic scale, examining the stress fields and the strains related to the contact between two solids. The Hertz theory for a normal contact has been established in 1881 [Hertz, 1881]. It allows the determination of the contact area and the stresses and strains distributions from the contact pressure, the solids geometry and their elastic moduli. In the following, only a brief introduction to the Hertz contact theory is made. For additional details, refer to books from Johnson [Johnson, 1985] or François et al. [François, 1991]. The problem for two convex solids is described by the Figure V.9.

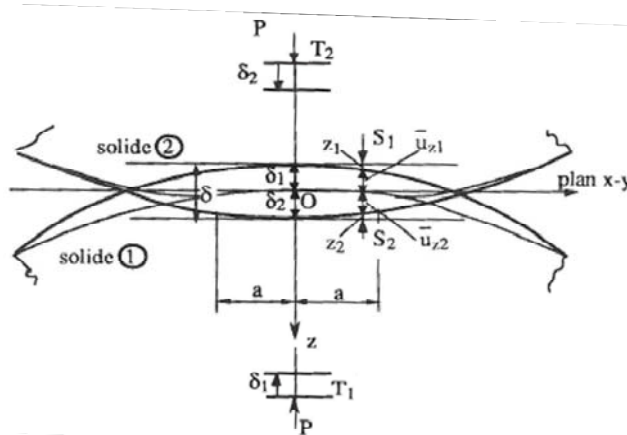


Figure V.9. Hertz contact for two convex solids, after [François, 1991].

The assumptions for an elastic contact are: (i) the surfaces are continuous, (ii) small strains are considered, (iii) each solid can be considered as a semi infinite solid and (iv) no sliding between the solids. It means that the contact radius (a) is very small compared to the curvature radius (R) of the two solids ($a \ll R$), and the friction coefficient is null ($\mu=0$).

For a bi-dimensional contact between two cylinders, the pressure distribution is parabolic, similarly to solids of revolution:

$$p(x) = \frac{2P}{\pi a} \sqrt{1 - \left(\frac{x}{a}\right)^2} \quad (\text{V.2})$$

Then, the contact radius and the maximum of the pressure are given by:

$$a = 2\sqrt{\frac{PR}{\pi E^*}} \quad (V.3)$$

$$p_0 = \frac{2P}{\pi a} = \sqrt{\frac{PE^*}{\pi R}} \quad (V.4)$$

where P is the load applied, E^* and R are respectively the equivalent Young's modulus and the equivalent curvature radius, defined by: $1/E^* = 1/E_1 + 1/E_2$ and $1/R = 1/R_1 + 1/R_2$, and p_0 is the pressure.

The various stresses undergone by the solid for such a contact are depicted in the Figure V.10. The following stresses are obtained:

$$\sigma_{xx} = \frac{-p_0}{a} \left((a^2 + 2y^2) * (a^2 + y^2)^{-1/2} - 2y \right) \quad (V.5)$$

$$\sigma_{yy} = -p_0 a (a^2 + y^2)^{-1/2} \quad (V.6)$$

$$\tau_1 = 1/2 |\sigma_{zz} - \sigma_{xx}| \quad (V.7)$$

The maximum shear stress τ_1 is equal to $0.30 * p_0$ for a distance $0.78 * a$. For solids of revolution, this maximum shear stress is found maximum at a distance $0.48 * a$.

In our case, the probe is sliding over the surface, and the die is immobile during the test. Thus, the friction force, which restrains the relative motion of the two solids, has to be considered. It modifies the contact stress. Additionally, a Coulomb friction law is used. The variation of the stress distribution due to the sliding is shown in Figure V.11.

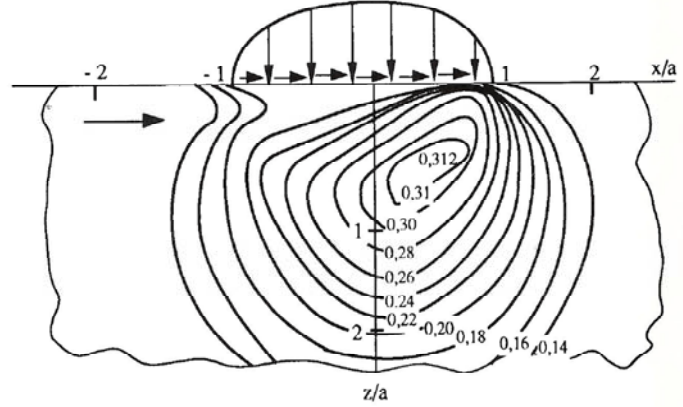
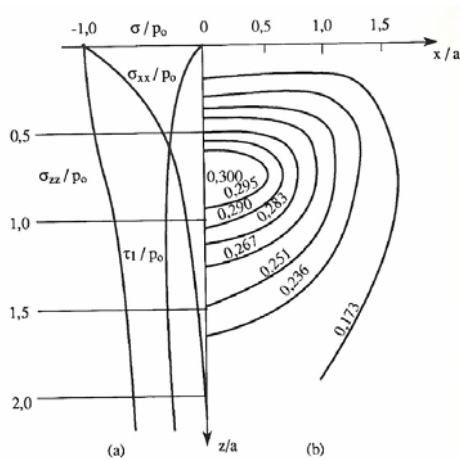


Figure V.10. Contact between two cylinders, stress in the structure (a) and iso-values shear stresses (b), after [François, 1991].

Figure V.11. Shear stress (τ_1) level curves in the structure for a sliding contact, with $\mu=0.2$, after [François, 1991].

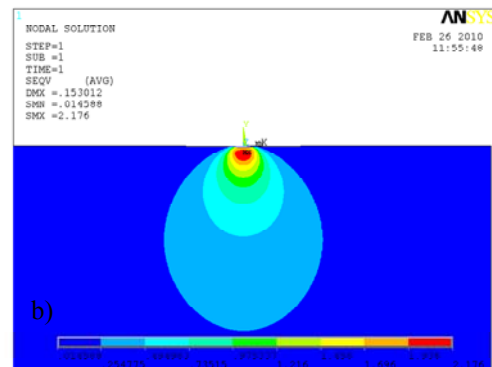
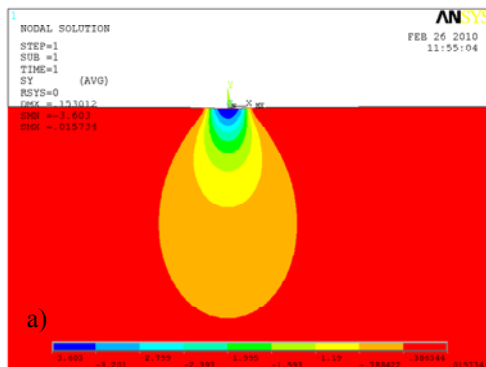


Figure V.12. Compressive (a) and Von Mises (b) stresses in silicon ($a=2\mu\text{m}$).

The Hertz contact is simulated in the silicon, using the pressure distribution aforementioned in the equation (V.2). The contact radius is 2 μ m. The stresses in the structure are depicted in the Figure V.12 and the Figure V.13. Similar results than those observed on the Figure V.10 are obtained. Hence, the Hertz contact is fairly reproduced from modelling in only one layer.

II.B.3.b) Analytical equations for the loading evaluation

These analytical equations must depend on all the test parameters and apparatuses (needle geometry, probe card...) to reproduce accurately the loading conditions. Hence, due to results availability and since these are the extreme loading cases, only needle geometries level 1 and level 4 are later considered. These analytical equations are derived from the calibration measurements described in the section III.B of the chapter 2.

For the whole range of technology nodes and probe card, the analytical equation describing the vertical (and also the lateral) reaction force is:

$$F_{vert} = \alpha_0 + \alpha_{11} * x_1 + \alpha_{12} * x_1^2 + \alpha_{21} * x_2 + \alpha_{22} * x_2^2 + \alpha_3 * x_1 * x_2 \quad (V. 8)$$

where x_1 is the overdrive (in μ m) and x_2 the number of passes, and the α_i are the fitting coefficients, given in the Table V.3.

Due to the important number of reaction forces measured by the PC provider Technoprobe, a second order equation is possible. Indeed, the vertical and lateral reaction forces have been evaluated for three overdrives (OD = 50 μ m, 75 μ m and 100 μ m) and for six different numbers of passes (1, 2, 3, 5, 7 and 9). Thus, the analytical equations are valid in the range of values considered during the experiments.

Technology and PC	CMOS 120nm and PC 1				CMOS 45nm and PC3			
Needle	L1		L4		L1		L4	
Reaction force	F _{vert}	F _{lat}	F _{vert}	F _{lat}	F _{vert}	F _{lat}	F _{vert}	F _{lat}
α_0	-0.02743	-0.03059	0.12296	0.04908	2.4677	1.1436	3.6543	1.7229
α_{11}	0.04044	0.00295	0.036	0.00098	n/a	n/a	n/a	n/a
α_{12}	0.00005	0.00001	0.00013	0.00003	n/a	n/a	n/a	n/a
α_{21}	0.00710	0.02184	0.03327	0.02473	0.0882	0.235	0.0294	0.0362
α_{22}	-0.00076	-0.00167	-0.00345	-0.00251	-0.0071	-0.0186	-0.002	-0.0024
α_3	0.00009	0.00016	0.00027	0.00049	n/a	n/a	n/a	n/a

PC and needle	PC 2, level 1			PC2, level 4		
Reaction force	F _{vert}	F _{lat}		F _{vert}	F _{lat}	
Technology	All	C 120nm	C 45nm	All	C 120nm	C 45nm
α_0	-0.03813	0.12601	1.5704	0.70875	0.57202	1.57
α_{11}	0.03387	-0.00163	n/a	0.02631	-0.01222	n/a
α_{12}	n/a	0.00004	n/a	n/a	0.0001	n/a
α_{21}	0.03625	-0.0105	0.2115	0.02313	0.01717	0.1625
α_{22}	-0.00313	0.00071	-0.0194	-0.00438	-0.00159	-0.0125
α_3	0.00013	0.00011	n/a	0.00044	0.00004	n/a

Table V.3. Values of the fitting coefficients in the analytical equations.

For the fitting constant in the analytical equation, the r-square is always found above 0.99. The standardized Pareto charts for the vertical and lateral reaction forces are shown in the Figure V.13 and Figure V.14. As demonstrated in the chapter 2, the overdrive is by far the

main contributor to the variations of the vertical reaction force. The lateral reaction force is also influenced in a more important range by the number of passes.

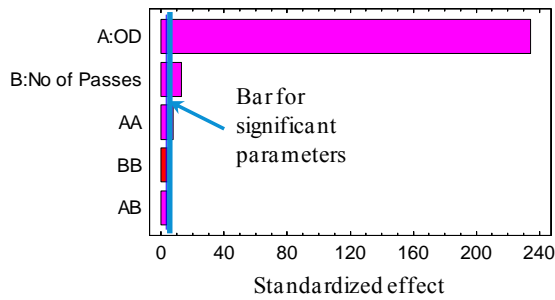


Figure V.13. Standardized Pareto chart for the vertical reaction force (PC 1, needle geometry 4).

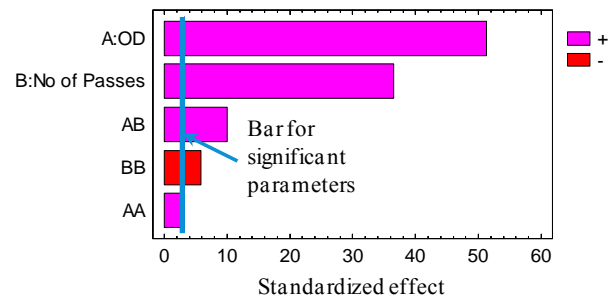


Figure V.14. Standardized Pareto chart for the lateral reaction force (PC 1, needle geometry 4).

Thus, by using these analytical equations, the vertical and lateral reaction forces undergone by the pad can be estimated for a wide range of test parameters and also technology nodes and probe cards and needle geometries. The pressure applied to the model is calculated from this abacus. As a consequence, all the parameters mentioned need to be inputs of the numerical model.

II.B.3.c) Application of the loading conditions

In the previous two sections, the Hertz contact and the vertical and lateral reaction forces abacus have been described, and are now used as bases of the loading. Initially, the target was to reproduce the whole probing test conditions in a single loading, however, due to the assumption of the Hertz contact (i.e. the maximum stress location dependence on the contact radius) and particularly the large contact area (long scrub marks), the maximum Von Mises stress was located in the silicon, which is not correlated to the cracks location in the pad open layer. Thus, multiple loading steps are applied to the pad. Indeed, the multi-step loading is suitable due to the probe kinematics, as further explained.

As shown in the section I.C of the chapter 4, the probe needle deflects during the test, and at the end of the loading, there is an angle between the pad surface and the bottom of the tip, which induces the high stresses obtained during the simulations. Then, the SEM observation of the cracks has shown that the crack is located in the middle of the probe mark. Hence, the contact of the probe on the pad is probably close to the one described in the Figure V.15. Thus, the contact area between the probe and the pad probably decreases during the loading stage. On the contrary, the vertical and lateral reaction forces increases.

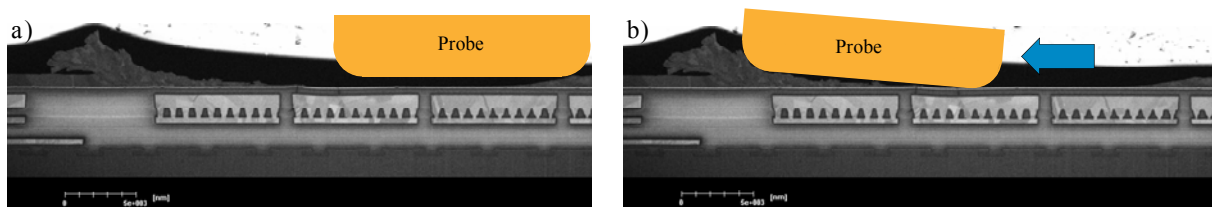


Figure V.15. Schematic representation of the initial contact (a) and contact assumed at the end of the loading (b).

Finally, the reaction force, the multiple loading steps and the variation of the contact area need to be introduced in the model. Thus, a Hertz contact with decreasing the contact area is used to reproduce the loading conditions, as shown by the Figure V.16. The contact area decreases from the tip radius ($8\mu\text{m}$) to the tip curvature radius (on the edges of the needle tip; $2\mu\text{m}$). The pressure (P) applied to the surface's elements in the location i is:

$$P(i) = \frac{j}{nb_step} \frac{F_{vert}}{\pi a^2} \sqrt{1 - \left(\frac{i - centre}{a} \right)^2} \quad (V. 9)$$

where:

- j/nb_step is the ratio (between 0 and 1) describing the increase of the vertical reaction force along the probe path, i.e., the ratio is null at the impact time and then reaches 1 for the last load step, when the vertical reaction force applied is maximum. This linear increase of the ratio is related to the linear correlation between the overdrive applied, the lateral displacement of the probe and the vertical reaction force measured, as previously described in the section III.B in the chapter 2.
- F_{vert} is the vertical reaction force, a is the contact area, and $centre$ is the location of the centre of the pressure application where the pressure is maximum for the given loading step.

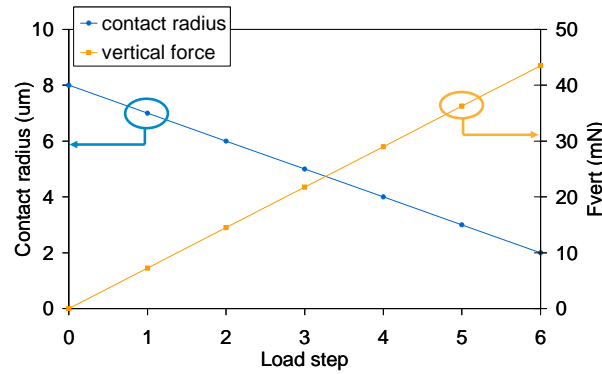


Figure V.16. Evolution of the contact radius and the vertical reaction force for various load steps.

Using such successive loading conditions implies a sharper and increasing pressure applied to the pad along the probe path (that goes from right to left), as shown in Figure V.17. This formulation allows obtaining a maximum load located before the end of the loading area, and to reproduce the increasing stress undergone by the structure.

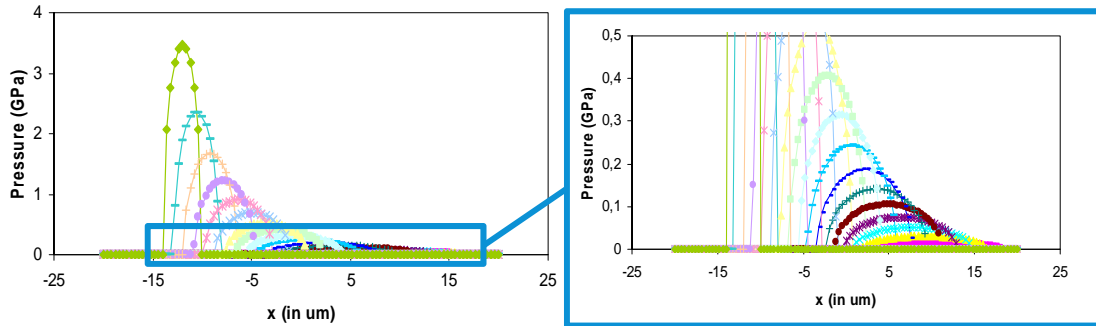


Figure V.17. Pressures applied to the pad during the successive load steps.

Similarly, tangential loading conditions are applied, using the lateral reaction force instead of the vertical one in the equation (V. 9).

II.B.4. Results and limitations of the 2D FE model

II.B.4.a) Stress analysis in the structure

The Von Mises stress in the structure for the successive load steps is shown in the Figure V.18. The successive load steps allow the introduction of the time in the finite element model

to reproduce the whole EWS loading stage. For these simulations, the OD is fixed to $90\mu\text{m}$, and thus the scrub length is about $40\mu\text{m}$, as observed in the section I.B.1 of the chapter 4. For the calculation of the vertical reaction force, the probe card 1, and the needle geometry 1 are selected. The first load step, which corresponds to the impact of the probe on the pad, is not depicted in the figures since the vertical reaction force is null. Then, a seven steps loading condition is applied to the pad. Obviously, the higher the number of steps, the closer to the actual loading condition the test is. However, this parameter increases the CPU time consumption, and consequently the duration of one simulation.

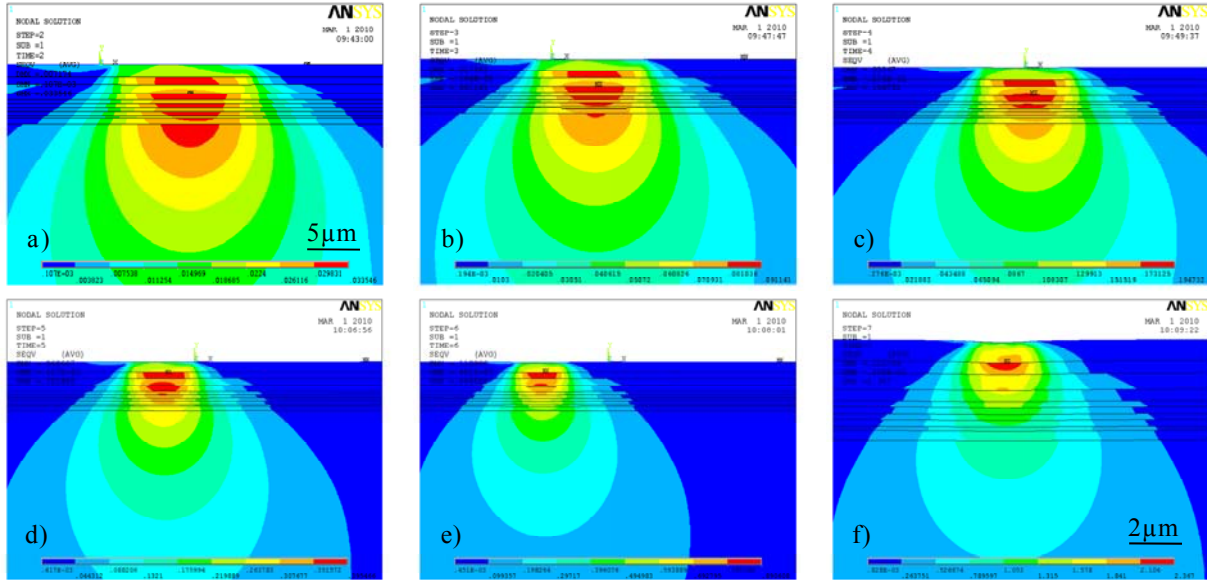


Figure V.18. Von Mises stress in the structure during the successive load steps (mind the range of stresses on each view).

The Von Mises stress in the pad increases during the successive loading steps, which is expected due to the increasing loading conditions. The evolution of the maximum stress in the pad is depicted in the Figure V.19. The variation of the location of the highest stress undergone is also shown. It is found that the highest Von Mises stress is initially located in the IMD3 layer and then, while decreasing the contact area, it moves to the PO layer, which is the layer within major cracks are observed.

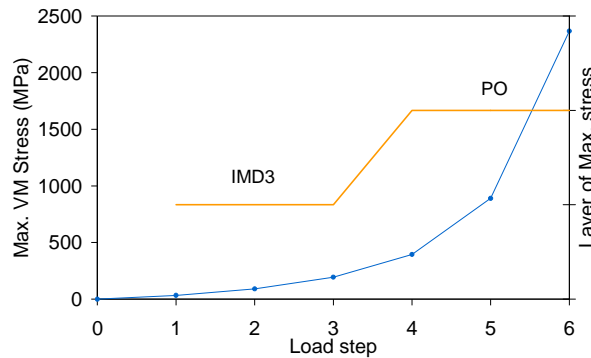


Figure V.19. Maximum Von Mises stress in the pad structure, and related location, as a function of the successive load steps.

Finally, the 2D model and the applied loading conditions are able to reproduce the EWS. The stresses undergone by the pad are still very high, but once more, the use of an almost punctual contact is the root cause of this estimation. Moreover, these values are found to be of

the same magnitude (2GPa to 4GPa) than those of the explicit 3D model that strengthen the loading assumptions performed.

II.B.4.b) Discussion

The effect of applying the Hertz contact on a multi-layer system is shown to modify the location of the stress peak in the structure due to the various material properties of the layers. Additionally, the effect of the location of the maximum of stress undergone by the pad is closely linked to the contact radius. Indeed, the Figure V.20 shows the variation of the Von Mises stress along the vertical axis (y-axis) in the centre of the pressure application ($i=0$ in equation (V. 9)). For the multi-layer system with quite high contact radius ($8\mu\text{m}$), the maximum is found in the oxide IMD3 (for $y/a=0.37$, that is $-2.96\mu\text{m}$) layer, as previously described. On the contrary, while only the silicon is considered, the stress peak is found deeper into the surface (for $y/a=0.67$).

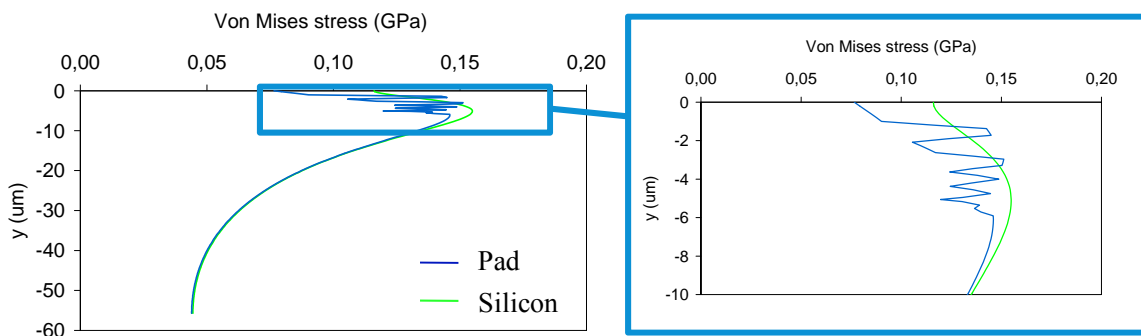


Figure V.20. Von Mises stress in the pad over the silicon substrate and in the silicon only along the vertical axis (for $i=0$, and a Hertz pressure centred in that location). The contact radius is fixed to $8\mu\text{m}$.

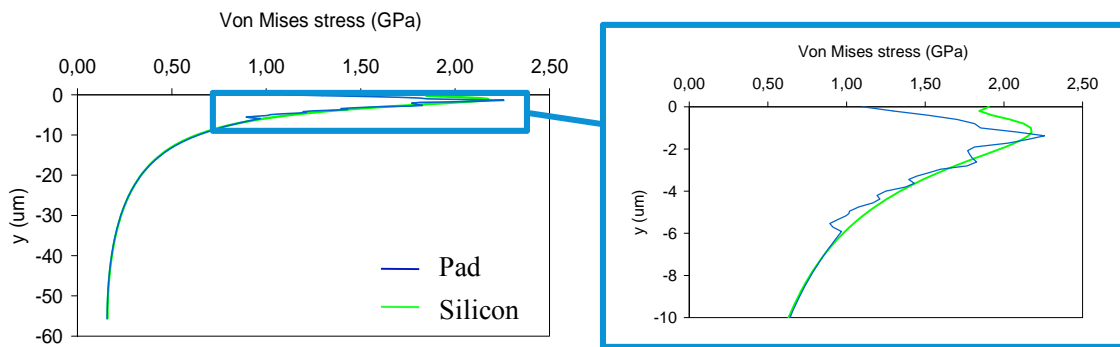


Figure V.21. Von Mises stress in the pad over the silicon substrate and in the silicon only along the vertical axis. The contact radius is fixed to $2\mu\text{m}$.

While the contact radius is lower ($a=2\mu\text{m}$), the maximum Von Mises stress is found in the pad open oxide layer. Then, the location difference with the actual stress peak described by the Hertz contact theory is smaller, as depicted in the Figure V.21. For such contact, the location is $y/a=0.6$. Hence, the value found is in the range of values given in the section II.A.3.a.

The behaviour at the bottom of the pad is not affected by the multi-layer system on top of it, whatever the contact radius, indicating a reduced effect of the bottom boundary conditions on the simulated results in the pad.

Some limitations due to the loading conditions and the use of a 2D model are underlined. As a matter of fact, a 2D model cannot describe entirely a 3D pad design. Indeed, layouts with

vias or CB vias, slotted copper plates and in a more general manner all the layouts that include 3D variations and orthotropic material properties within a layer cannot be evaluated through a 2D model.

Concerning the loading conditions, the results for the study of crossed parameters (e.g. variation of several layer thicknesses, layouts...) must be considered with care. In this model, only elastic material properties have been used, and reproducing the scrub mark requires to use elastic-plastic material properties for aluminium and to remove the elements, after a given threshold penetration is reached. Similarly, the work hardening of the aluminium is not taken into account. Finally, due to the contact assumptions, and the stress simulated in the structure, the comparison between the structures must remain qualitative. Yet, the trends from these qualitative analyses to compare the mechanical robustness of the pads is expected to be true since a deep study of the loading conditions have been performed.

II.C. Development of the 3D implicit FE model

II.C.1. Description of the 3D FE model of the pad

The 3D geometry of the pad is similar to the explicit 3D model's one described in the chapter 2. Hence it consists in a half pad due to the symmetry axis of the centre of the probe mark. Additionally, some silicon is added to reduce the boundary effect and the mesh is much smaller, allowing the introduction of small details. The pad is meshed with SOLID186 elements, whose size is fixed at $0.5\mu\text{m} \times 0.5\mu\text{m} \times 0.2\mu\text{m}$, respectively in the plane and for their thickness. The meshed pad is shown in the Figure V.22. For the top layers, some quadrangle elements are selected, and the silicon is meshed with tetrahedrons, with bigger sizes, in order to reduce the number of elements in this layer.

To reproduce the loading conditions, some SURF154 elements are used. These are extensions of the elements used in the 2D model to the 3D model. The same material properties are used as the 2D model (see Table V.2).

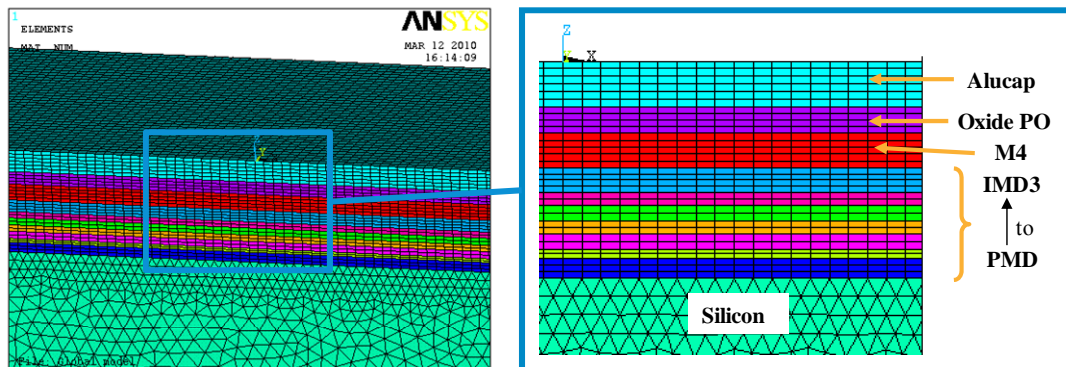


Figure V.22. 3D meshed model of the pad.

The same boundary conditions as the 2D model are applied (for the nodes of the lateral and bottom of the pad). Moreover, the node on the symmetry axis ($y=0$) and for the highest y values have their DOF along the y -axis fixed.

II.C.2. Loading conditions

Along the symmetry axis of the model, the loading conditions are the same as those used in the 2D model, i.e. based on the Hertz contact and described in the section II.B.3, with the same profile (see Figure V.23 a). The problem concerns the third dimension and the reproduction of the probe mark along its width. Initially, a linear decrease of the pressure along the y -direction is applied, as shown in the Figure V.23 (b). The purpose is to reproduce

as fairly as possible the loading induced by the EWS. Hence, the pressure applied to the surface's elements for the last simulation step is depicted in the Figure V.23 (c).

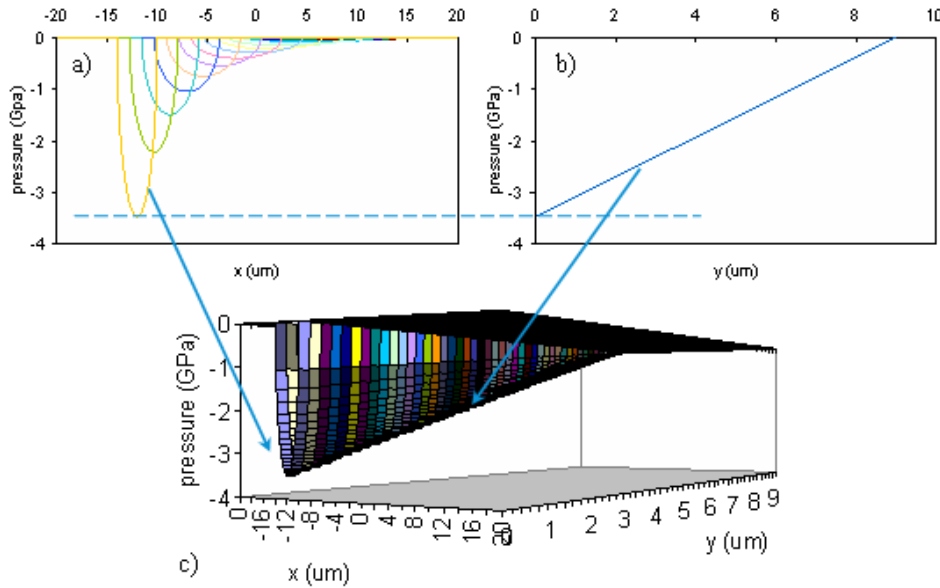


Figure V.23. Loading conditions: for the successive load steps along the symmetry axis (a), along the third dimension, the y-axis (b) and 3D representation of the x-y load for the last load step (c).

II.C.3. Results of the 3D model

II.C.3.a) Stress distribution in the pad

The compressive stress along the y axis in the aluminium is firstly investigated. The compressive and Von Mises stresses for the last step are depicted in the Figure V.24 (a) and (b). This loading step is considered due to the higher stress induced and the use of elastic material properties in the pad. As expected, both stresses decrease linearly on the y-axis. At the end of the scrub width, the compressive stress is null, and the Von Mises one is low (few tens of MPa).

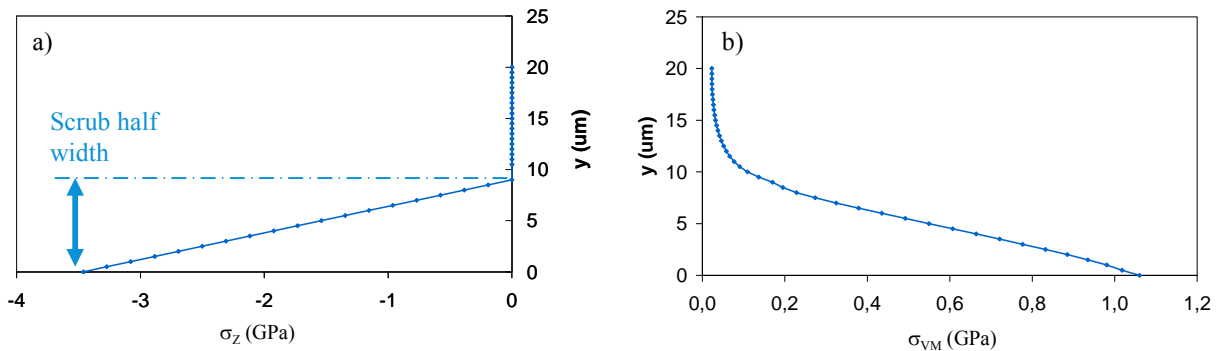


Figure V.24. Compressive stress (σ_z) applied along the y-axis in the aluminium ($z=0$) (a), and corresponding Von Mises stress (b).

Moreover, the Von Mises stress in the various oxide layers of the pad is shown in the Figure V.25. The maximum of the compressive stress is found in the aluminium, as shown in the Figure V.26 a). The maximum Von Mises stress in the structure is found in the pad open oxide layer, similarly to the 2D model (see Figure V.26 a).

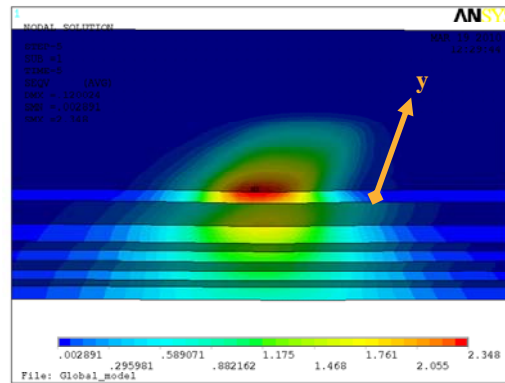


Figure V.25. View of the Von Mises stress in the oxide layers of the pad.

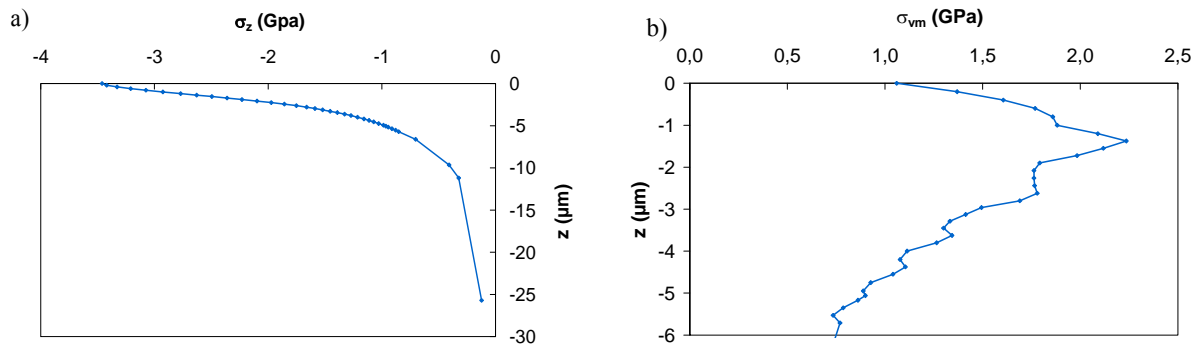


Figure V.26. Compressive stress applied along the z-axis (a), and corresponding Von Mises stress (b).

II.C.3.b) Evaluation of the stacking effect

In the chapter 4, increasing the thicknesses of the aluminium and the pad open oxide layers have been shown to induce a reduction of the failure hazard. It has been obtained by means of experimental tests (i.e. cracks occurrence) and numerical modelling (using stress analysis). Thus, in this section, the extreme stacks are weighted, i.e. thinnest thicknesses of alucap and PO versus the thickest ones.

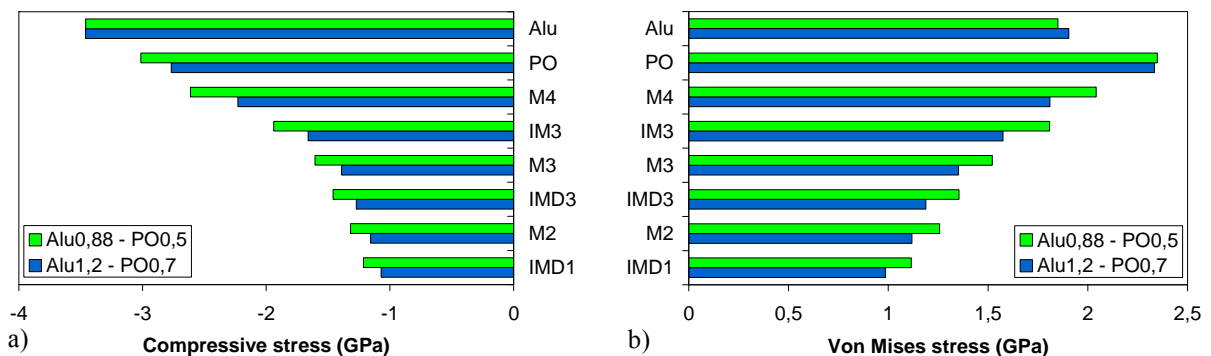


Figure V.27. Maximum compressive (a) and Von Mises stress in the various layers of the pad for the thinnest and thickest pad stacks.

The Figure V.27 shows the maximum stresses in the various layers of the pad. Increasing the thickness of the aluminium and pad open are shown to reduce the stresses in the PO and IMD3 layers by respectively 0.9% and 14.5%. Yet, low stress variations in the layers are observed since elastic isotropic material properties are used for the layers (and particularly the copper ones, where the design has been shown to have an effect on the stress in the oxide).

Addition of details within the pad such as vias implies stress concentration points and as a consequence higher variations between the stacks.

On the Figure V.27 (b), slightly higher maximum stress is observed in the aluminium layer for the thickest stack. It is explained by the Hertz contact profile in the pad. Indeed, the stress peak is obtained in the bottom of the aluminium layer. Hence, this result is different from the simulations made using the 3D explicit model (see Figure IV.25, in the previous chapter). However, during the EWS test, the aluminium undergoes high plastic deformation which is not reproduced in the two numerical models. In the section I.B.3 in the chapter 4, the failures obtained with the Berkovich tip during the nanoindentation have shown that the fracture presumably starts in the top of the oxide layer.

II.C.4. Insertion of an elastic-plastic behaviour for the aluminium

II.C.4.a) Selection of the elastic-plastic law and values

The EWS and the nanoindentation have underlined the plastic deformation of the aluminium capping the pad, and its role as a stress buffer for the oxide layers underneath. Thus, elastic-plastic material properties are introduced in the 3D implicit model.

For a perfectly elastic-plastic solid, the Tabor relationship [Tabor, 1948] express that the hardness of a material under an indenter is a constant factor of the yield stress:

$$H = 3 * \sigma_y \quad (V. 10)$$

The hardness of the aluminium has been measured by means of CSM nanoindentation and found to be close to 0.6GPa. Thus, the yield stress of the aluminium is set to 200MPa.

The choice of the tangent modulus (describing the plastic regime) is arbitrarily made. Indeed, the stress strain curves, and consequently the tangent modulus defining the plastic regime for such thin films depend on the process, the thickness... Thus, in the numerical model, it is either fixed at 6GPa or at 40GPa. These tangent moduli are selected to have a trade-off between noticeable plastic deformation and the convergence issues in the numerical model. The lowest value is selected to slightly increase the plastic deformation. Finally, the stress-strain curve for the aluminium is depicted in the Figure V.28 for the lowest tangent modulus.

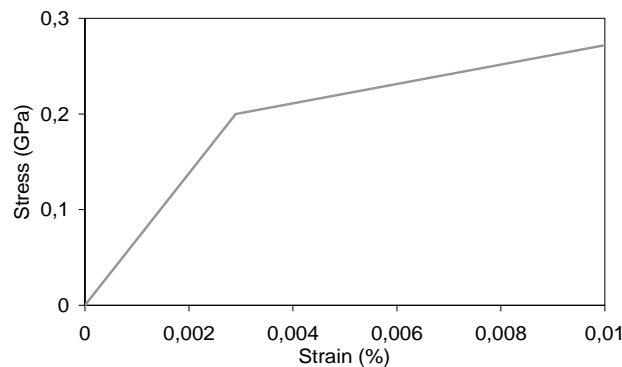


Figure V.28. Stress-strain curve for the aluminium layer (plastic tangent modulus equal to 6GPa).

Then, in the numerical model, two plastic options can be selected [Ansys, rel. 11], i.e. the Bilinear ISOTropic hardening (BISO) and the Bilinear KINematic hardening (BKIN). The BISO option assumes an isotropic work hardening, which is produced by a homogeneous multiplication of the number of dislocations. The BKIN assumes that the total stress range is equal to twice the yield strength, so that the Bauschinger effect is included. This effect is related to a heterogeneous distribution of the dislocations, and so the internal stress resulting.

Due to the various assumptions (particularly the stress range) and the best results obtained for large deformations, the bilinear isotropic behaviour is selected.

The use of elastic-plastic properties increases the CPU time consumption. Indeed, the duration to solve the model with 10 loading steps goes from about 4 hours to about 20 hours. It is explained by the necessity to introduce substeps for the various loading steps to ensure the convergence of the numerical model.

As a matter of fact, the introduction of plasticity into the numerical model allows checking through comparison with actual scrub mark, the reliability of the loading conditions. Besides, the effect of the number of passes can finally be investigated.

II.C.4.b) Determination of the lateral profile

In the section II.C.3.a), a linear decrease of the stress along the y-axis has been applied. The use of elastic plastic material properties for the aluminium allows the observation of its plastic deformation and the possibility to reproduce the shape of the scrub mark. Then, three different profiles along the y-axis are selected:

	for $0 \leq y \leq \text{scrub width}$	for $y > \text{scrub width}$
- Linear:	$p(y) = p_0 \left(1 - \frac{y}{\text{scrub_width}} \right)$	0
- Parabolic:	$p(y) = p_0 \left(1 - \left(\frac{y}{\text{scrub_width}} \right)^2 \right)$	0
- Plateau:	$p(y) = p_0$	0

These loading profiles of pressure are depicted in the Figure V.29. The scrub width is set to $9\mu\text{m}$, which is about the half width for the scrub mark for a $90\mu\text{m}$ overdrive (see section I.B.1 in the chapter 3). The compressive stress resulting is shown in the Figure V.30.

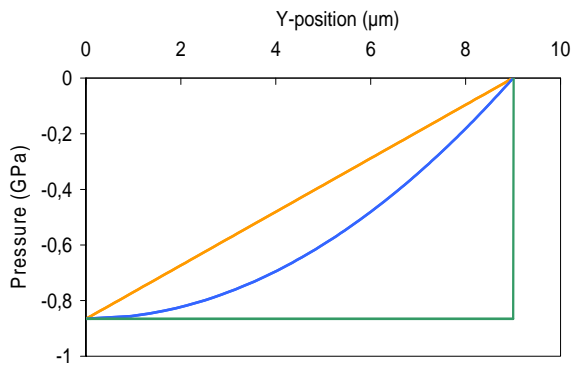


Figure V.29. Pressure profiles along the y-axis for the last loading step.

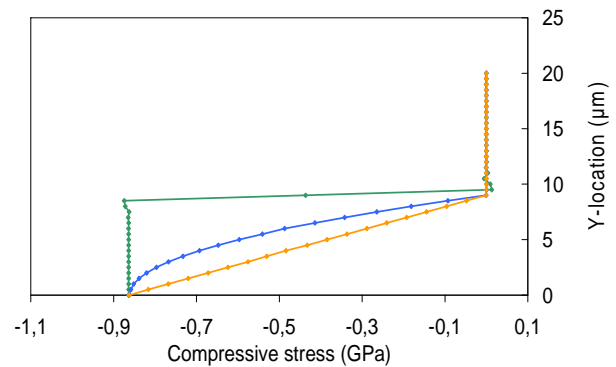


Figure V.30. Compressive stress in the aluminium for the various loading profiles.

The top view of the equivalent cumulated plastic strain in the aluminium for the linear profile is shown in the Figure V.31. The plastic deformation is cumulated over the 6 successive loading conditions applied to the pad to reproduce the EWS. The plastic deformation in the lateral direction does not exceed 40% of the scrub width. Hence, this profile does not allow a fair reproduction of the mark. Similar view of the plateau loading profile is depicted in the Figure V.32 a). The plastic deformation of the aluminium does not

match the observation of an actual scrub mark. Moreover, stress peaks are observed in all the layers due to the sharp variations of the loading condition at the probe mark edge. Thus, this loading profile does not reproduce the scrub mark.

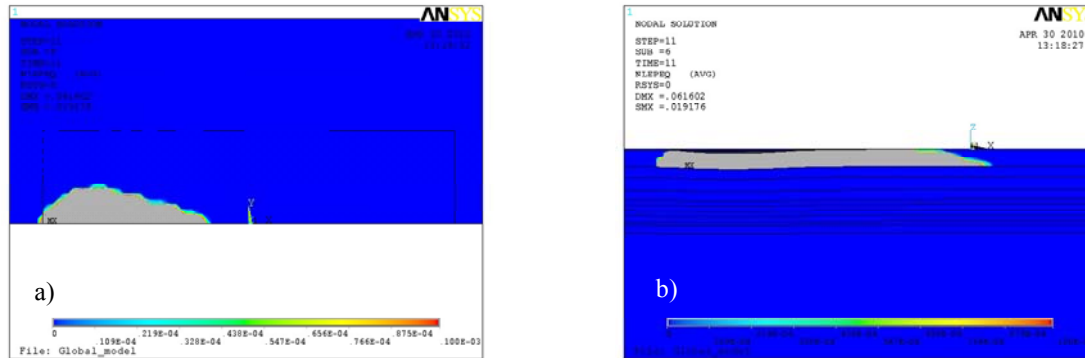


Figure V.31. Top view of the equivalent cumulated plastic strain in the aluminium for the linear loading profile (a) and cross-section view of the equivalent plastic strain in the pad.

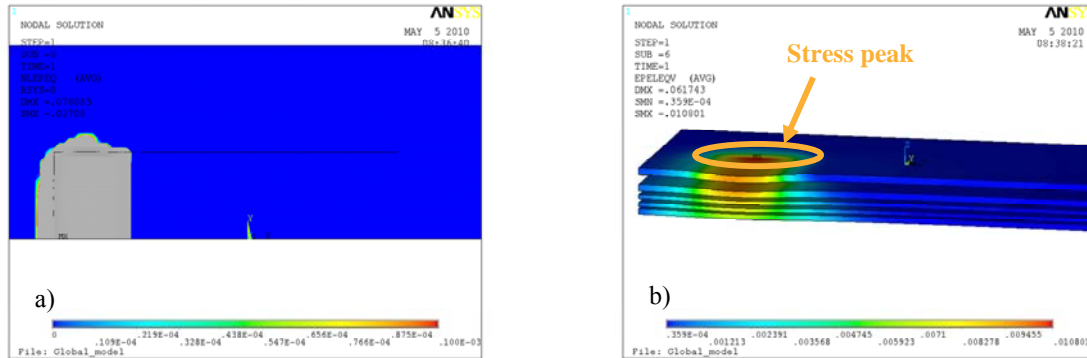


Figure V.32. Top view of the equivalent cumulated plastic strain in the aluminium for the plateau loading profile (a) and Von Mises stress in the oxide layers for the same profile (b).

Then, the cumulated equivalent plastic strain in the aluminium for the parabolic loading profile is investigated, as shown in the Figure V.33. The plastic deformation in the y-direction reaches about 70% on the top of the layer. Moreover, this value reaches almost 100% while performing similar observation on the bottom of the aluminium layer, as depicted by the Figure V.33 b). Wider plastic deformation is observed at the bottom of the aluminium. It has been previously demonstrated that the maximum stress in the aluminium is located at the bottom of the layer, and then the threshold to plastically deform is more rapidly obtained in that location.

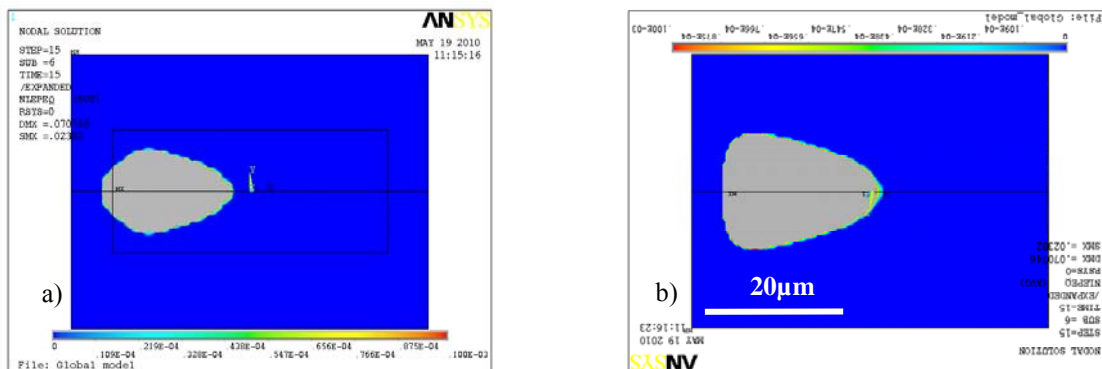


Figure V.33. Cumulated equivalent plastic strain: top (a) and bottom (b) views of the aluminium.

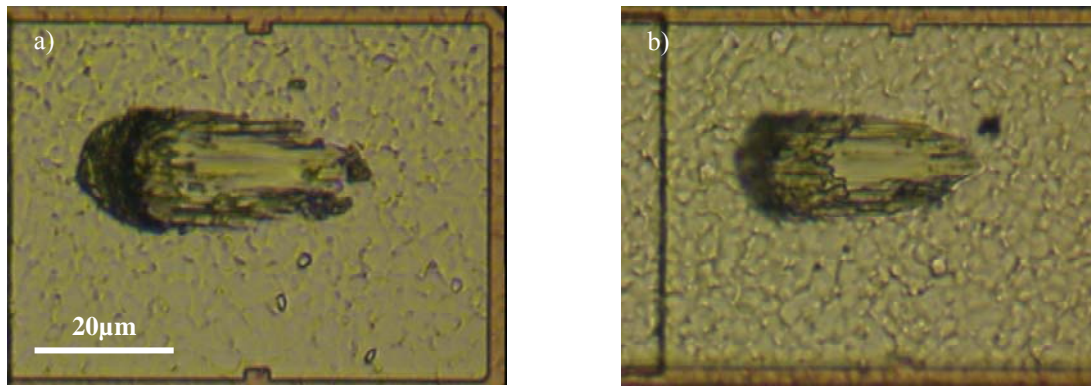


Figure V.34. Optical microscopy top views of scrub marks in the pad, for OD 90μm, after 5 passes (a) and 10 passes (b).

Actual probe marks are shown in the Figure V.34. The comparison between the plastic deformation (particularly the top view) in the simulated model and the actual ones shows a quite good reproduction.

The actual probe mark is longer than the simulated one (Figure V.33). It is explained by the low pressures applied to the pad in the numerical model during the first loading steps. Modifying the linear increase of the vertical reaction force and decrease of the contact radius can reduce this effect. Yet, for the EWS tests, it can mean that at the beginning of the test (i.e. at the impact location), the aluminium is not plastically deformed. In the literature, Strom [Strom, 1998] has observed similar results and shown that the probe mark can be small compared to the probe path for extreme cases, as shown in Figure V.35. Nevertheless, in-situ observation of the EWS test is complex for the cantilever probe card and needle geometry used. However, this assumption is probably correct.

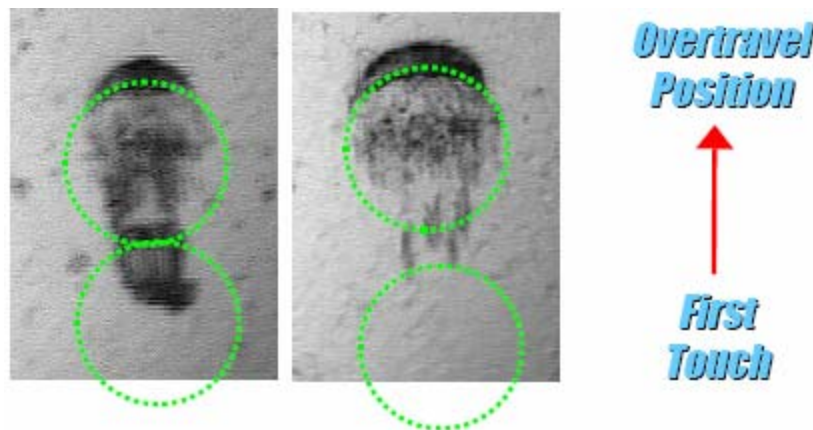


Figure V.35. First touch and overtravel positions and resulting probe mark for two different needles, after [Strom, 1998].

II.C.4.c) Effect of the EWS tests parameters

The effect of the overdrive and the number of passes are evaluated with the 3D implicit FE model and elastic-plastic material properties. More specifically, their effect on the stress in the structure and the plastic deformation of the aluminium, i.e. the probe mark size, is investigated.

Much longer probe marks have been observed with the optical microscope in the section I.B.1 in the chapter 3. The Figure V.36 shows the cumulated equivalent plastic strain in the aluminium. As expected, the plastic deformation is longer and wider while the OD increases.

The maximum plastic strain in the aluminium layer and the maximum Von Mises stress in the oxide layer increases as well.

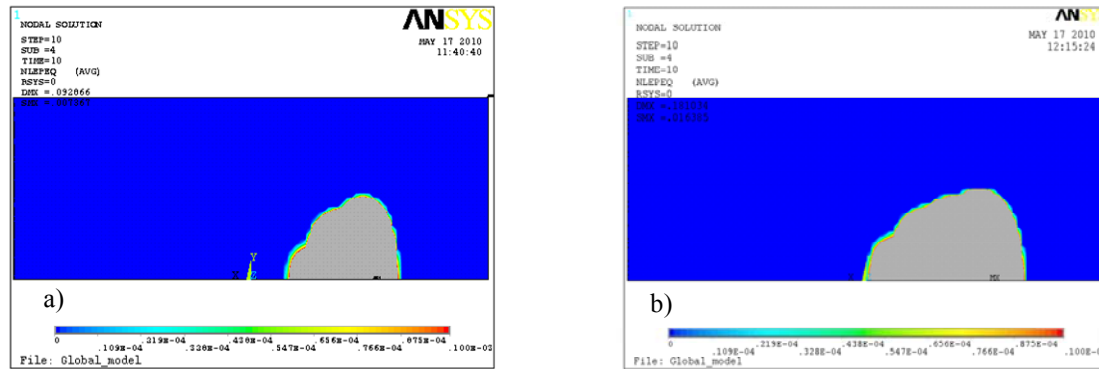


Figure V.36. Bottom view of the cumulated equivalent plastic strain in the aluminium, for overdrives of 50μm (a) and 90μm (b).

On the effect of the number of passes, due to an important computation time, only probing tests with one and two passes are evaluated. The second pass induces a small increase of the Von Mises stress in the oxide (2.532GPa, +1.4%) and a small increase of the plastic deformation of the aluminium ($16.79 \cdot 10^{-3}$, +2.4%). These small augmentations are related to the small variations of the measured vertical reaction force, while increasing the number of passes (see section II.A.3.b)).

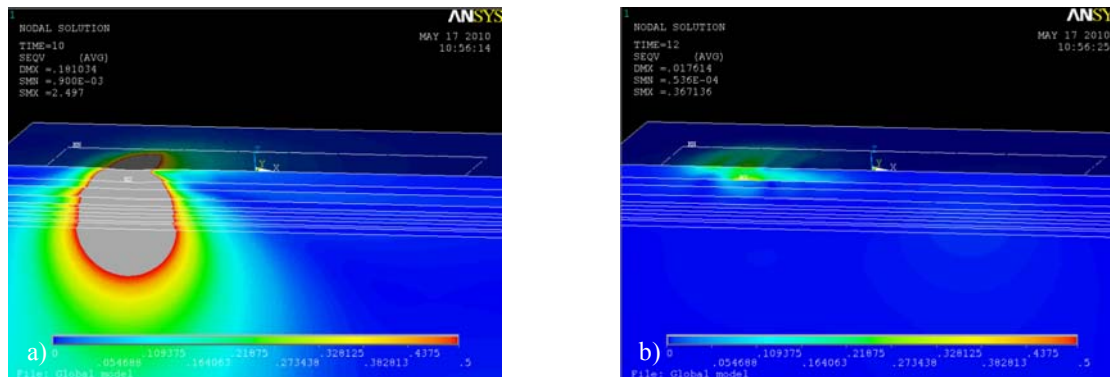


Figure V.37. Von Mises stress on the structure at the last step of the first pass (a) and at the first step of the second pass (b).

Some compressive residual stress in the aluminium is observed after the first pass ($\sigma_{VM}=367\text{MPa}$), as depicted in the Figure V.37 b). Residual stress is also observed in the pad open oxide layer.

Finally, the introduction of elastic-plastic material properties for the aluminium in the numerical model allowed an enhanced determination of the probe mark, and more specifically of the stress and strains of the lateral profile. Moreover, the same trends as actual probe mark measurements have been observed while considering the plastic deformation of the pad capping layer. Despite the good reproduction of the probe mark, some additional fine tuning of the aluminium properties (tangent modulus), and of the loading profile (linear decrease of the contact radius) can further improve it. However, the use of elastic-plastic material properties implies a non negligible increase of the CPU time per simulation. Then, the for large pad studies, the 3D model with the elastic material properties is trustable and sufficient to evaluate and rank various pad layouts.

II.D.EWS simulations: explicit versus implicit 3D modelling

In this section the explicit (using ANSYS LSDyna, see previous chapters) and the implicit (using ANSYS multiphysics, current model) 3D finite element models are compared, focussing on the expectations for the evaluation of the test and pad layouts parameters.

		Explicit 3D model	Implicit 3D model
EWS test parameters	Needle dynamics	Very efficient	Not available
	Chuck velocity and vibrations	Very efficient	Not available
	Overdrive	Very efficient	Very efficient
	Number of passes	Moderate: high CPU consumption, only through variations of the reaction forces	Moderate: use of elastic-plastic properties but no modelling of the contact
	Other test parameters (general expectations)	Efficient	Moderate due to the loading conditions and the absence of contact
Pad layouts	Effect of layer thickness (alucap, pad open...)	Very efficient	Very efficient
	Effect of a given design (e.g. presence of via)	Moderate: good results but limited due to the use of homogenized properties	Very efficient
	Pad layout (e.g. lines vs. slots)	Moderate: same as previously	Very efficient
	Other pad layout parameters (general expectations)	Moderate due the non representation of small details	Very efficient: Model developed specifically for the pad layout comparison, no size limitation

Table V.4. FEM ability to evaluate the test parameters and pad layouts.

The Table V.4 summarizes all the capabilities of the two numerical models. In a general manner, the explicit model is very efficient to evaluate all the test parameters due to the reproduction of the contact between the pad and the probe needle. Then, for the analysis of pad layouts, this model is efficient while large details and variations are investigated. While studies of small details are required, the use of the homogenization technique can reduce the investigation field. On the contrary, the implicit model is not very efficient to evaluate the tests parameters because the probe is not introduced in the model. However, it is dedicated to the study of pad layouts, and the very small mesh allows the introduction and the evaluation of very small details. Besides, the CPU time consumption difference between the two models has to be considered for large designs of experiments. Actually, the simulation of the explicit model lasts about four days, while this of the implicit one only takes a few hours.

Finally, depending on the requirements of the study, both finite element models are very useful and efficient to evaluate the probing related issues in the pad.

III. Conclusions

In this chapter, both experimental and numerical efficient industrial tools are proposed for the optimization of the pad structure. For the nanoindentation, the comparison between the Berkovich and the cube corner tip is initially realized on the tip's geometrical point of view. Thus, the Berkovich tip is shown to induce larger indent for a similar maximum displacement into the structure. Then, the effect of the choice of tip on the load-displacement curve is

considered. The cube corner tip is seen to induce high stresses and strains near the tip, and exhibits a local behaviour of the pad. However, it also implies a larger dispersion of the measurements than the Berkovich tip. After that, the material properties of the copper deduced from the indentation with both tips is described, and lower values of the Young's modulus and hardness are obtained with the cube corner tip due to the Oliver and Pharr model limitation (pile-up not taken into account). Finally, general guidelines for the pad's study are given and also some indications about the tip selection while indentation with both tip cannot be performed.

In a second part, a three dimensional implicit finite element model is setup to reproduce the EWS. This model is based on all the experimental and numerical information previously collected to define relevant loading conditions. Initially, a 2D model is developed to test the ability of successive Hertz type loading conditions to reproduce the overall probing test. Once the loading conditions of the two dimensional model have shown their ability to reproduce the EWS ones, this model is extended to a 3D finite element model. Then, elastic-plastic material properties are introduced for the aluminium layer, to check the ability of the model to reproduce the probe mark. Hence, an enhanced loading profile is selected. The effect of the overdrive and the number of passes is shown to increase the plastically deformed aluminium area, i.e. the scrub mark, similarly to actual tests. These results on the elastic-plastic model have proven the ability of the loading conditions applied to reproduce the EWS. As a conclusion, the 3D explicit and implicit finite element models are compared and the advantages and drawbacks of each of them are underlined in the frame of the probing test and pad studies.

References

[Ansys, rel. 11] Ansys user manual, release 11

[Ayhan, 1999] Ayhan, A.O. and Nied, H.F. , Finite element analysis of interface cracking in semiconductor packages, IEEE Transactions on components and packaging technology, vol. 22, no. 4, December 1999.

[Bucaille, 2003] Bucaille, J.L., Strauss, S., Felder, E., Michler, J., Determination of plastic properties of metals by instrumented indentation using different sharp indenters, Acta materialia, vol. 51, 2003, pp. 1663-1678.

[Bucaille, 2004-A] Bucaille, J.L., Rossoll, A., Moser, B. Strauss, S. and Michler, J., Determination of the matrix in situ flow stress of a continuous fibre reinforced metal matrix composite using instrumented indentation, Materials science and engineering A, vol. 369, 2004, pp. 82-92.

[Bucaille, 2004-B] Bucaille, J.L., Strauss, S., Schwaller, P. and Michler, J., A new technique to determine the elastoplastic properties of thin metallic films using sharp indenters, Thin solid films, vol. 447-448, 2004, pp. 239-245.

[Chollacoop, 2003] Chollacoop, N. Dao, M. and Suresh, S., Depth-sensing instrumented indentation with dual sharp indenters, Acta materialia, vol. 51, 2003, pp. 3713-3729.

[Durst, 2006] Durst, K., Backes, B., Franke, O. and Göken, M., Indentation size effect in metallic materials: Modelling strength from pop-in to macroscopic hardness using geometrically necessary dislocations, Acta Materialia, vol. 54, 2006, pp. 2547-2555.

[François, 1991] François, D., Pineau, A. and Zaoui, A., Comportement mécanique des matériaux, Hermes, 1991.

[Hertz, 1881] Hertz, H., Über die berührung fester elastischer körper, Journal für die reine and angewandte mathematic 92, 1881, pp. 156-171.

[Jang, 2008] Jang, J.I., Pharr, G.M., Influence of indenter angle on cracking in Si and Ge during nanoindentation, Acta materialia, vol. 56, 2008, pp. 4458-4469.

[Johnson, 1985] Johnson, K.L., Contact mechanics, Cambridge university press, 1985.

[Johnson, 2007] Johnson, M., Li, Z., Wang, J. and Yan, Y., Mechanical characterization of zeolite low dielectric constant thin films by nanoindentation, Thin solid films, vol. 515, 2007, pp. 3164-3170.

[Krueger, 2002] Krueger, R., Paris, I. L., O'Brien, T.K. and Minguet, P.J., Comparison of 2D finite element modeling assumptions with results from 3D analysis for composite skin-stiffener debonding, Composite Structures, vol. 57, 2002, pp. 161-168.

[Lagace, 2002] Lagace, P.A., Plane stress and plane strain, unit 6, Massachusetts institute of technology, 16-20 fall 2002, available in <http://www.core.org.cn/NR/rdonlyres/Aeronautics-and-Astronautics/16-20Structural-MechanicsFall2002/02E38870-4C45-4732-AD80-B0FE55FA6B0D/0/unit6.pdf>, consulted on September, 29, 2010.

[Mirshams, 2004] Mirshams, R.A. and Parakala, P., Nanoindentation of nanocrystalline Ni with geometrically different indenters, Material science and engineering A, vol. 372, 2004, pp. 252-260.

[Mirshams, 2006] Mirshams, R.A. and Pothapragada, R.M., Correlation of nanoindentation measurements of nickel made using geometrically different indenter tips, *Acta Materialia*, vol. 54, 2006, pp. 1123-1134.

[Nix, 1998] Nix, W.D. and Gao, H., Indentation size effect in crystalline materials: a law for strain gradient plasticity, *Journal of the mechanics and physics of solids*, vol. 46, no. 3, 1998, pp. 411-425.

[Strom, 1998] Strom, J., Multi tier probe cards and contact resistance, Presentation at the IEEE Semiconductor wafer test workshop, 1998, <http://www.swtest.org/archive.html>, consulted on September, 29, 2010.

[Tabor, 1948] Tabor, D.A., A simple theory of static and dynamic hardness, *Proceeding from research society*, London, ser. A 192, 1948, pp. 247.

[Wei, 2006] Wei, P.J. and Lin, J.F., Determination for elasticity and plasticity from time dependant-nanoindentation, *Material science and engineering A*, vol. 496, 2008, pp. 90-97.

[Wu, 2006] Wu, F., Gonchond, J.-P., Optimisation de la technique de nanoindentation pour la caractérisation des propriétés mécaniques des oxydes basse permittivité, Université Joseph Fourier, Grenoble, France, Master degree thesis, 2006.

[Yeo, 2004] Yeo, A., Lee, C. and Pang, K.H.L., Flip chip solder joint fatigue analysis using 2D and 3D FE models, *IEEE international conference on thermal, mechanical and multiphysics simulation and experiments in micro-electronics and micro-systems*, 2004, pp.549-555.

Conclusion and perspectives

I. Synthesis

The first chapter initially describes the main stages of the manufacturing of an integrated circuit. A focus is made on the pad, which corresponds to an opening in the chip to realize the electrical connections with the outer part before the packaging operations. In the semiconductor industry, the current trend to reduce the critical dimensions and introduce low-k, brittle dielectric materials, threatens the integrity of the bond pad and failures in the oxide layers of the interconnections are observed.

On pads, the electrical wafer sort (EWS), or probing, is performed to check the electrical functionality of the die at the wafer level. During this test, a contact between a stiff needle and the aluminium on top of the pad is established. As a consequence, the aluminium undergoes high damages and the integrity of the layers underneath is a concern. Then, the various assemblies that enable the electrical connections with the package are presented. A particular focus is made on the wire bonding, which is the most common way to achieve the connection. It consists in bonding the pad to the outer part of the I.C. with gold or copper wires. The wire pull test and the ball shear test allow the characterization of the wire adhesion and the intermetallic compounds created. The failure modes observed during these tests can reveal a mechanically weak structure.

In the second chapter, the various experimental techniques to characterize and gain insight on the failures in the pad such as the FIB milling followed by SEM observations or the optical profilometry are introduced. Afterwards, an experimental method, the nanoindentation, is used to reproduce the EWS loading conditions and compare the mechanical robustness of the pads. This experimental technique is set up to reproduce the loading conditions and failures related to EWS and assembly steps. The Berkovich and cube corner tips, used during testing are also presented.

Then, the explicit finite element model to mimic EWS is described, starting with the description of the geometries of the needles investigated and the geometrical model corresponding to the pad. The physical (material properties) and numerical (meshing, effective volume...) parameters used in the model are justified. In order to build the model and to fairly reproduce the probing, the calibration is described and a good agreement is obtained between the experimental dedicated measurements and the model. Finally, a modal analysis is performed to evaluate the influence of the numerical (mass scaling) and physical parameters (sliding).

In the third chapter, the damages induced by the test and process parameters are considered. The objective is to induce failures (cracks during EWS or peeling during wire pull tests) in the structures to discriminate the parameters, even if these must be avoided during the manufacturing of an integrated circuit. In that frame, the current inertial effects during the probing (velocity and vibrations of the chuck) are initially shown by means of the evaluation of the simulated reaction forces to unlikely induce the cracks in the pad. Then, pad integrity

analyses on dies and the numerical model indicate that the increase of test parameters (the overdrive from 50 μm to 90 μm and the number of passes from 5 to 10) is detrimental to the aluminium capping layer and also to the pad beneath.

Then, several bonding process parameters (bonding power, time and force and contact force) are evaluated through designs of experiments. It is demonstrated for the gold bonding that the increase of the parameters altogether induces an increase of the diameter and a decrease of the height of the welded ball. The shear tests provide insights concerning the effect of the bonding parameters on the gold ball adhesion to the pad. Due to the current trend to move toward copper bonding for process overall cost reduction, similar study is performed for such wire. Thus, the contact force and the bonding power are questioned. The increase of the bonding power (from 70mW to 90mW) and the reduction of the contact force (from 50mN to 30mN) reduce the peeling occurrence in the structure. Hence, the contact area between the pad and the welded ball is highlighted to be a key parameter to avoid fractures in the interconnections. In parallel, the shear tests showed that contrary to gold bonding, the increase of the bond strength between the copper ball and the aluminium pad is often related to an increase of the peeling risks. Moreover, in this chapter, guidelines to control test and process parameters and reduce the fracture hazard in the pad structure are proven. Finally, it is demonstrated that the process and test parameters show higher preponderance than the layout ones to imply failures in the pad. However, during the die manufacturing, it is often not possible to modify the parameters aforementioned, which are defined with respect to a product application.

In the fourth chapter, most of the currently available designs are considered through experiments and modelling in order to strengthen the pad robustness. Initially, many EWS tests, wire pull and shear tests and related experimental characterizations are performed to gain insight on the failure mechanisms. These tests also provide a baseline for the ranking of the structures according to their mechanical robustness for further comparison with nanoindentation and simulations. Cracks and peeling cases are located in the pad open oxide or within the inter-metal dielectric layers. More precisely, the geometrical singularities in the pad (for a given layout) are pointed out to be detrimental to the pad robustness. On the contrary, the protective behaviour of aluminium and copper during these tests and their role on the mechanical robustness of the pad is underlined. Then, the finite element model and the nanoindentation are used to mimic these test and process. The stress analysis performed on the results from the numerical model provides reliable trends for the structure optimization. Similarly, for the nanoindentation, the maximum displacement into surface and a novel indicator, which allows the determination of the critical load for the crack appearance in the oxide layer, are successfully employed to reproduce the actual tests results.

Concerning the pad designs effects, all the techniques show that increasing the thicknesses of the aluminium (from 0.88 μm to 1.2 μm) and pad open oxide (from 0.5 μm to 0.7 μm) layers is related to a reduction of the failure occurrence in the pad open oxide and inter-metal dielectric layers. Then, the metal density of the copper layer is demonstrated to have an effect on the mechanical robustness of the pad. However, the copper density variation is intimately linked to the insulating materials surrounding it (presence of silicon glass or low-k). The presence of the pad open oxide layer is observed to improve the resistance to compression. However, while vias are inserted within the layer, under the probing or the bonding areas, the numerous geometrical singularities introduced as well increase dramatically the number of peeling cases. The layout of the top copper level can be modified in order to reduce the crack occurrence. Indeed, the pad layout with slotted plates, and particularly with smallest oxide holes, should be preferred to wide copper lines while designing the pad. The location of the plates with respect to the probing location is also shown to reduce the crack occurrence.

Finally most of the pad layout parameters investigated have an effect on the mechanical robustness of the pad, and their modification according to the results previously described helps to reduce the fracture hazard. Additionally, a full set of tools is given to compare the pad architecture. The use of nanoindentation and finite element modelling is able to reproduce most of the probing and some of the wire bonding issues. Hence, these methods can be used for technology development and to solve crises by giving reliable alternative to the expensive EWS and wire bonding operations applied to the pad.

In the last chapter, general nanoindentation guidelines for the evaluation of the mechanical robustness of pads are given and an implicit finite element model dedicated to the evaluation and enhancement of the pad structures is set up. Indeed, in an industrial context, fast and reliable prototyping are required. In that frame, the Berkovich and the cube corner tip are initially compared from a geometrical point of view. Thus, the Berkovich tip is shown to induce larger indent for a similar indentation depth. On the contrary, the cube corner tip induces higher stresses and strains at the vicinity of the tip, and exhibits a local behaviour of the pad. Then, the effect of the tip choice on the load-displacement curve and for the comparison of pad structures is discussed. Finally, rules for the study of pads and for the tip selection are given.

In a second part, an implicit finite element model is setup to reproduce the EWS. This model is based on all the experimental and numerical works previously achieved to define relevant loading conditions. Elastic-plastic material properties are introduced for the aluminium layer, to check the ability of the model to reproduce the probe mark. Hence, an adequate loading profile is deduced. The effect of the overdrive and the number of passes is shown to increase the plastically deformed aluminium area, i.e. the scrub mark, similarly to actual tests. These results on the elastic-plastic model prove the ability of the loading conditions applied to reproduce the EWS. Finally, the 3D explicit and implicit finite element models are compared and the advantages and drawbacks of each of them are underlined in the frame of the control of the probing test and the optimization of pad architectures.

II. Perspectives

The perspectives of this work are multiple. Concerning the study of pads, the correlations have been established between the experimental and numerical techniques. However, the difficulty to determine a critical load prior to failure has been underlined. This is due to the complexity of the multilayer system composing the pad structure and to the lack of tools at the sub-micrometre scale for such characterization. These thresholds prior to failure can be used during modelling to rank the structures and define critical loading conditions. In addition, some deeper understanding of the failure initiation and propagation can provide insights with respect to the enhancement of the pad structure. It can be achieved through the development of in-situ techniques during the test and process.

Concerning the reproduction of the test and process, nanoindentation has shown some limitations to reproduce mechanical issues induced by the wire bonding. Therefore, additional techniques such as scratch test after shallow indentation can provide complementary information. Indeed, the small lateral displacements could reproduce the shear stresses undergone by the pad during the application of the ultrasonic power during the bonding process. In that frame, the wire pull and shear tests, which induce additional tensile and/or shear stresses in the structures are also questioned for their efficiency to characterize the effect of the bonding process itself. The development of additional standards for the bond characterization to evaluate the damages due to process only and obtain information such as

its induced residual stress in the structure can further allow the optimization of the process parameters and the bond pad architecture. Thus, a characterization such as the removal of the ball, the aluminium and the barrier layers by chemical etches similarly to the pad integrity test can provide some of the information (on cracks in the interconnections).

Then, the forward step of the semiconductor industry to reduce the critical dimensions requires the continuous development of new finite element models to follow the evolution of the manufacturing changes. Indeed, numerical model describing the EWS test with cantilever probe card has been successfully set up. Yet, in the next few years, the pad pitch diminution will probably imply to move toward probing with vertical probe cards. Even if the root causes for the failure hazard are expected to be similar to these observed for current probe cards, the ratio between the vertical and the lateral reaction forces radically changes, requiring the development of new explicit models and corresponding calibration to gain insight on the tests.

In addition, finite element models, either using implicit or explicit schemes are also required to evaluate the assembly processes, obviously the wire bonding but also the flip chip and the three dimensional integration that also occurs on the pads and raise some pad integrity concerns. For the wire bonding the numerical model would consist in fact in a least two models: one for the bonding process itself, and one or more for the wire pull and ball shear tests, within residual stresses due to the wire bonding process are introduced. These numerical models can allow the reproduction of the results obtained from the actual tests and the reduction of the failure occurrence in the interconnections.

Appendixes

Outline

A.	Comparison between nanoindentation, EWS and wire bonding	196 -
A.1.	Pad layout description	196 -
A.2.	EWS results	197 -
A.3.	Nanoindentation tests	197 -
A.3.1.	Results with the Berkovich tip	197 -
A.3.2.	Results with the cube corner tip	197 -
A.4.	Wire bonding	198 -
A.5.	Correlations between techniques	198 -
B.	DOE used for the evaluation of the test and assembly parameters and the pad layout designs	199 -
C.	Acronyms	200 -

A. Comparison between nanoindentation, EWS and wire bonding

During the development of new applications of pads on the CMOS 180nm technology node, some mechanical issues have been observed in the pads with buchettes. As a consequence, a design of experiment with several pad architectures is performed. Thus, nanoindentation with Berkovich and cube corner tip is performed in parallel to actual EWS test and bonding operations (probing and bonding performed by ST team).

A.1. Pad layout description

For such a technology node, only aluminium is used for the metal levels, whether located on top of the structure or inside (aluminium is replaced by copper in the latest technology nodes), as depicted in the Figure 1. During the DOE, five different pad architectures are considered to solve the mechanical issues observed after EWS. Within the DOE, the effect of the metal density and the design of the M2 aluminium layer and the presence of vias between the layers are investigated. The Figure 2 shows the four distinct M2 layouts considered, with varying aluminium density from 0% to 94%.

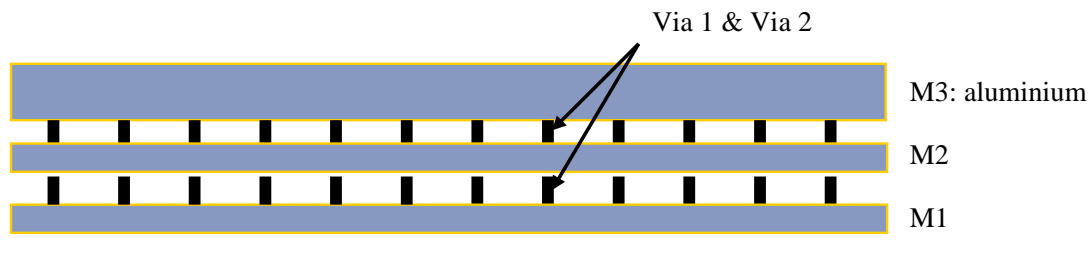


Figure 1. Schematic representation of the CMOS 180nm pad.

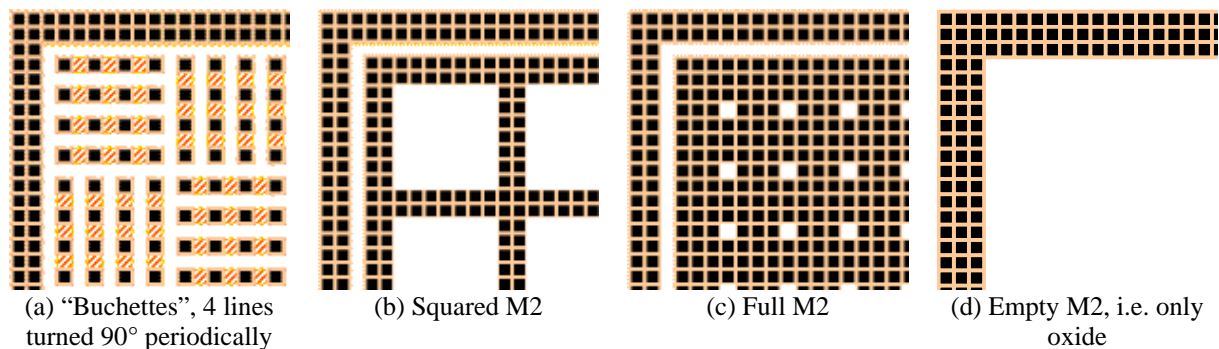


Figure 2. Schematic representation of the M2/V2 pad layouts (orange refers to aluminium, black refers to vias and white to oxide).

The various layouts investigated during the DOE are depicted in the Table 1. The standard configuration that initially raises questions is the configuration 1, within the second aluminium level is made from "buchettes".

Configuration	M2 layout	Presence of V1 and V2
1	Buches	Yes
2	Full	Yes
3	Full	No
4	Squared	No
5	Empty	No

Table 1. Pad layouts investigated during the DOE.

A.2. EWS results

Electrical wafer sort have been performed on these pads. The electrical response can then be evaluated. While mechanical issues occur in the pad, short circuits imply a null electrical response. Hence from the measurements, the pads 1, 2 and 3 depict some failures through all the performed trails. On the contrary, no electrical issue is detected for the pad structures 4 and 5. Some pad integrity tests are performed on the pads, more specifically, 10 dies on the edge of the wafer and 10 dies in its centre are randomly investigated. For the pad layouts 2 and 3, cracks are observed in most of the pads, and correlated to the electrical results. For the pads configurations 4 and 5, no crack is observed in accordance with the electrical measurements. For the pad option 1, no crack is observed during the inspection, but an inspection on the samples that failed to pass the electrical tests depicts some failures. Hence, a ranking according to the crack occurrences after EWS is established for these structures:

$$5 > 4 > 1 > 2 \approx 3 \quad (1)$$

The effect of the via is not shown to be significant. Moreover, the layout of the M2 level and more specifically the decrease of the aluminium density are shown to increase the mechanical robustness of the pads. Indeed, the latter effect is explained by the material properties of the aluminium and the oxide.

A.3. Nanoindentation tests

On each pad structure, 15 indentation tests are performed on the centre and the edge of the wafer for each layout, using the two tips.

A.3.1. Results with the Berkovich tip

The maximum load is fixed at 127.5mN. As described in the chapter 4, the maximum displacement into the surface is used to compare the resistance to the compression for each layout, as depicted on the Table 2.

Pad layout	Displacement at max. load in the centre of the wafer (nm)	Displacement at max. load at the edge of the wafer (nm)
1	1913 (23)	1864 (23)
2	1925 (22)	1878 (26)
3	1925 (40)	1881 (22)
4	1873 (23)	1823 (31)
5	1861 (25)	1788 (25)

Table 2. Maximum displacement and its standard deviation for the indentation with the Berkovich tip on the centre and at the edge of the wafer.

Hence, from the maximum displacement reach at the maximum load, a ranking according to the structures can be established:

$$5 > 4 > 1 > 2 \approx 3 \quad (2)$$

Yet, for the measurements with the Berkovich tip, in some cases, the standard deviation is higher than the shift between the curves. As a consequence, some measurements with the cube corner tip are complementary achieved.

A.3.2. Results with the cube corner tip

For the nanoindentation with the cube corner tip, the load is set to 14.7mN. No crack is observed in the oxide crack, probably due to the mechanical properties of the aluminium. Then, the results of the measurements are shown in the Figure 3 and summarized in the Table 3.

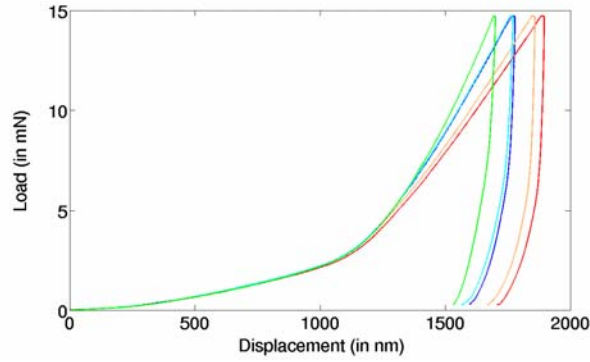


Figure 3. Averaged load versus displacement curves for the nanoindentation with the cube corner tip in the centre of the wafer.

Pad layout	Displacement at max. load in the centre of the wafer (nm)	Displacement at max. load at the edge of the wafer (nm)
1	1765 (29)	1706 (22)
2	1882 (38)	1848 (27)
3	1845 (35)	1825 (20)
4	1759 (24)	1768 (29)
5	1692 (17)	1684 (23)

Table 3. Maximum displacement and its standard deviation for the indentation with the cube corner tip on the centre and at the edge of the wafer.

Hence, from the maximum displacement reach at the maximum load, a ranking according to the structures can be established:

$$5 \gg 4 \approx 1 \gg 2 \approx 3 \quad (3)$$

A.4. Wire bonding

Some wire pull and shear tests have been performed by implied ST teams after gold wire bonding. These have been performed on the standard pad (layout 1), and on the optimized option from the EWS and nanoindentation results (layout 5). Thus, they obtained for both layouts neck breaks during pull tests (i.e. pass the test). However, a deeper investigation to observe cracks after bonding shows that a very low amount of the pads depicts some cracks for the option 5, while about half of the pads with layout 1 show have failures. The effect of the thermal aging is even more detrimental to the layout 1.

A.5. Correlations between techniques

The five pad layouts are investigated by means of actual EWS tests and nanoindentation. As shown in the chapter 4, nanoindentation with the two tips is a very efficient tool to reproduce the EWS loading conditions. Indeed, similar ranking according to the mechanical resistance of the structures is observed for the two tests. Moreover in that case, the nanoindentation results are in accordance with those of the wire bonding, at least for the comparison between the layouts 1 and 5.

B. DOE used for the evaluation of the test and assembly parameters and the pad layout designs

Technology node	chapter 2	Yes	Yes	CMOS 120nm - A
Experimental measurements for the probe calibration	chapter 3	Yes	/	CMOS 120nm - B
EWS test parameters: OD and # of passes	chapter 3	Yes	Yes	CMOS 45nm
Gold and copper wire bonding	chapter 4	/	Yes	
Effect of the aluminium and pad open oxide thicknesses	chapter 4	Yes	/	
Effect of the metal density, the presence of via and PO via	chapter 4	/	Yes	
Effect of the top copper layer design	chapter 4	Yes	/	
Loading conditions of the implicit models	chapter 5	Yes	/	

The table describes the use of the DOE performed for the investigation of the test and design parameters (from chapters 2 to 5).

C. Acronyms

BD: Ball Diameter (wire bonding parameter)
BE: Back End
BEoL: Back End of Line
BF: Bonding Force (wire bonding parameter)
BH: Ball Height (wire bonding parameter)
BOP: Bonding over passivation (presence of an oxide layer between aluminium and copper)
BP: Bonding Power (wire bonding parameter)
BT: Bonding Time (wire bonding parameter)
CC: Cube Corner tip (nanoindentation tip)
CF: Contact Force (wire bonding parameter)
CMOS: Complementary Metal Oxide Semiconductor
CPU: Central Processing Unit
CSM: Continuous Stiffness Measurement
DOE: Design Of Experiment
EBSD: Electron BackScatter Diffraction
EDX: Energy Dispersive X-ray spectroscopy
EFO: Electrical Flame Off
EWS: Electrical Wafer Sort
FAB: Free Air Ball
FE: Front End
FEM: Finite Element Model
FIB: Focussed Ion Beam
HAZ: Heat Affected Zone (wire bonding process)
HTS: High Temperature Storage
IC: integrated circuit
IMC: InterMetallic Compound
IMD: InterMetal Dielectric
MX: Metal level (X=1...7)
OD: OverDrive (EWS parameter)
PC: Probe Card (EWS)
PF: Pull Force
PO: Pad Open oxide (layer below the aluminium)
SEM: Scanning Electron Microscopy
SF: Shear Force
SS: Shear Strength
TD: TouchDown
TSV: Through Silicon Vias
USG: Undoped Silicon Glass
WB: Wire Bonding

Analyse expérimentale et numérique des défaillances mécaniques locales induites dans les interconnexions par les tests paramétriques et assemblages : optimisation des procédés et des architectures des plots de connexion

Résumé

Sommaire

I.	Description des tests et processus induisant des défaillances dans les pads	204 -
I.A.	Tri électrique des plaques (EWS).....	204 -
I.A.1.	Description du test.....	204 -
I.A.2.	Défaillances associées et méthodes d'observation.....	205 -
I.B.	Assemblage filaire.....	206 -
I.B.1.	Description du processus.....	206 -
I.B.2.	Méthodes de caractérisation	206 -
I.C.	Problèmes mécaniques rencontrés	207 -
II.	Présentation des méthodes expérimentales et numériques mises en œuvre.....	208 -
II.A.	Observation et reconstruction du pad en 3 dimensions.....	208 -
II.B.	Evaluation de la robustesse des pads : nanoindentation.....	208 -
II.C.	Modèle numérique reproduisant l' EWS.....	209 -
III.	Effet des paramètres de test et d'assemblage sur l'apparition des défaillances	211 -
III.A.	Paramètres critiques lié au test EWS.....	211 -
III.B.	Connexion électrique par fil de cuivre	212 -
IV.	Effet de l'architecture des pads sur robustesse mécanique	214 -
IV.A.	Effet de l'épaisseur des couches supérieures	214 -
IV.B.	Effets de la densité de cuivre et de la présence des vias	218 -
IV.C.	Effet de l'agencement de la couche de cuivre.....	221 -
V.	Adaptation au contexte industriel.....	223 -
V.A.	Méthodologie pour l'étude des architectures de pad.....	223 -
V.B.	Développement d'un nouveau modèle numérique reproduisant l'EWS	225 -
VI.	Conclusions	227 -
	Références bibliographiques	229 -

La tendance actuelle dans l'industrie des semi-conducteurs est à une réduction des dimensions, que ce soit au niveau de la taille des transistors ou de leur volume global (intégration tridimensionnelle). De plus, de nouveaux matériaux fragiles (oxydes faible permittivité) sont employés, et de là, la robustesse mécanique des architectures tend à diminuer. Dans cette thèse, on s'intéresse plus précisément aux plots de connexion (pads) qui permettent de relier les circuits intégrés au monde extérieur. Ces derniers se composent du volume des interconnexions situé sous cette zone non couverte par des couches de passivation, qui protègent aussi bien physiquement que chimiquement la puce (cf. Figure 1). Sur ces pads, des tests de vérification de fonctionnalité et des opérations d'assemblage, ou connexions entre la puce et le boîtier sont réalisés. Au cours de ces tests, des fissures apparaissent dans les interconnexions provoquant des courts circuits et de fait rendent la puce non fonctionnelle.

Les objectifs sont donc d'une part de mieux comprendre les causes d'apparition de ces défaillances, afin de définir des règles de dessin au niveau des architectures de pads. D'autre part, il s'agit de proposer des méthodes expérimentales et numériques simplifiées mais donnant des tendances fiables. Ces outils doivent permettre de reproduire ces tests et assemblages pour comparer diverses architectures de pad dans le cadre industriel du développement technologique ou lors de la résolution de crises.

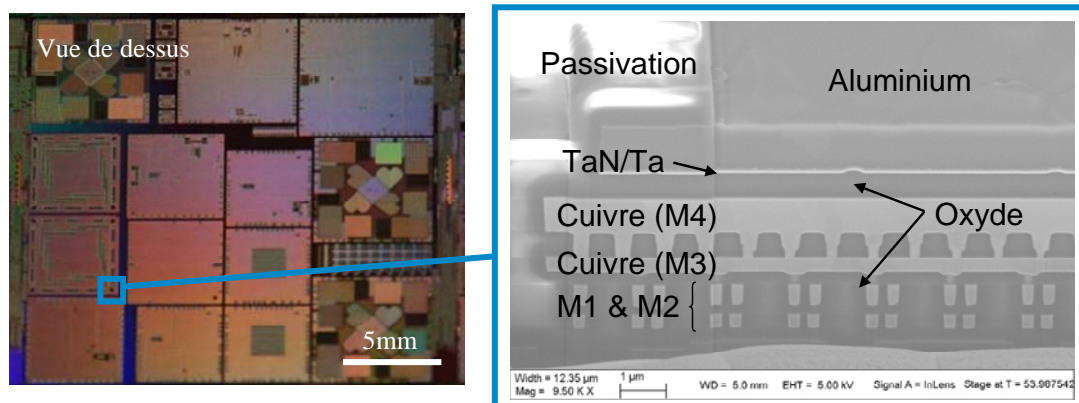


Figure 1. Vue de dessus de puces (à gauche) et vue en coupe (MEB avec un angle de 52°) d'un pad CMOS 120nm. Ce dernier contient 4 niveaux de cuivre (M1 à M4), avec un isolant entre ces couches (oxyde). La surface du pad se compose d'une couche d'aluminium.

I. Description des tests et processus induisant des défaillances dans les pads

Dans cette première partie, les tests de tri électrique des plaques (en anglais « Electrical Wafer Sort », EWS), ainsi que les processus d'assemblage des puces, et en particulier les connexions grâce à des fils d'or ou de cuivre (en anglais « Wire bonding ») sont décrites. Puis les défaillances induites dans les architectures de pad au cours de ces opérations sont introduites.

I.A. Tri électrique des plaques (EWS)

I.A.1. Description du test

Le test électrique des plaques est effectué à la fin des opérations de dépôt de couche sur le silicium et de réalisation des interconnexions entre les différents niveaux de cuivre. Il permet de vérifier la fonctionnalité électrique de la puce avant sa mise en boîtier. Ce tri est effectué depuis plusieurs dizaines d'années et consiste en établir un contact entre un élément de

contact (pointe, fils...) et le pad. Le résultat de ce frottement laisse une empreinte dans l'aluminium situé en surface du pad. A une échelle plus grande, un testeur, se composant d'une carte à pointe est utilisé pour faire les tests et distinguer les bonnes puces des mauvaises (voir Figure 2).

Le test se compose d'un cycle charge-décharge durant lequel une pause sert à effectuer la mesure électrique. Les paramètres importants au cours de ce test sont :

- le déplacement vertical de la plaque servant à déplacer les wafers,
- le nombre de passage nécessaire avant d'avoir une réponse électrique correcte de la puce,
- la vitesse de déplacement de la plaque automatisée.

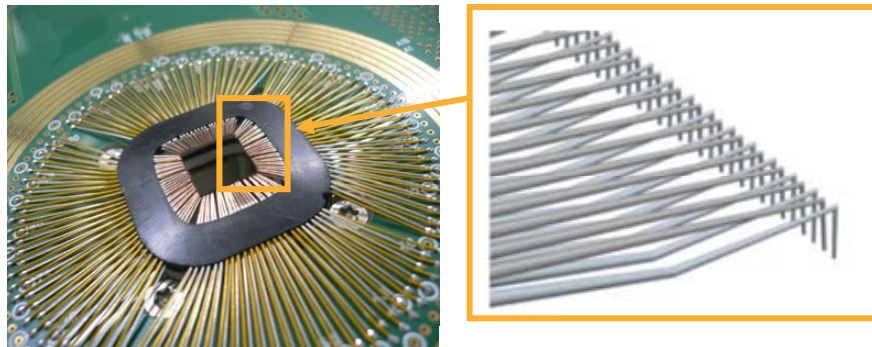


Figure 2. Carte à pointes cantilever (à gauche) et vues des pointes la composant (à droite), d'après [Vallauri, 2007].

I.A.2. Défaillances associées et méthodes d'observation

Au cours du tri des puces, l'aluminium en surface est sévèrement endommagé, comme le souligne la Figure 3. Des défaillances apparaissent également dans les niveaux d'interconnexion sous le pad. Actuellement, la méthode d'observation la plus répandue pour évaluer si une structure présente des défaillances ou non est un test d'intégrité du pad. Sur un échantillon testé, une solution chimique puis un plasma sont utilisés pour enlever les couches d'aluminium et de TaN/Ta en surface de la structure. La gravure chimique permet une visualisation de l'oxyde situé sous le pad. A la suite de cette gravure, une autre solution chimique s'infiltrant à travers les fissures dans l'oxyde provoque une gravure du cuivre, et permet une meilleure visualisation de ces défaillances sous microscope optique.

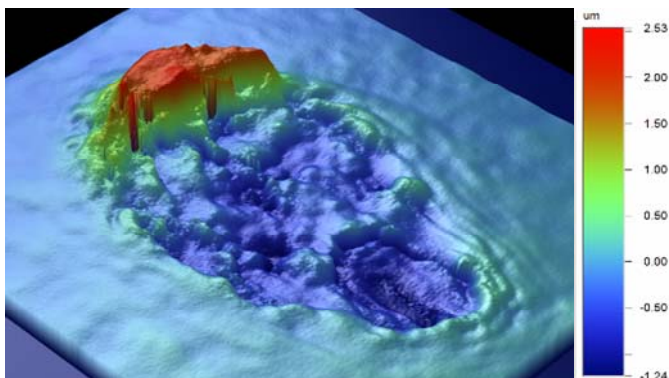


Figure 3. Observation en 3 dimensions de la surface d'un pad après EWS par profilométrie optique.

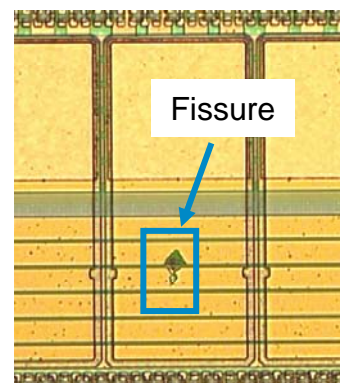


Figure 4. Révélation d'une fissure au cours du test d'intégrité du pad.

I.B. Assemblage filaire

I.B.1. Description du processus

La technique d'assemblage filaire (en anglais : « Wire bonding ») est la plus répandue en ce qui concerne la connexion entre le pad et le boîtier. Bien qu'ayant légèrement évolué au cours des dernières décennies, cette technique a été employée dès la fin des années 1950 [Anderson, 1957]. La technique actuellement employée est basée sur un soudage thermo-sonique en plusieurs étapes afin de souder une boule d'or ou de cuivre sur le pad en aluminium. Elle combine l'utilisation d'ultrasons et un chauffage des puces permettant de faciliter le soudage et entraîne la création de composés intermétalliques entre le matériau du fil et le pad. L'utilisation de l'or pour réaliser ces connexions est la plus répandue. Cependant pour des raisons économiques principalement, l'utilisation du cuivre se développe. Comparé à l'or il a des avantages tels que sa meilleure conductivité électrique et la plus faible vitesse de formation des composés intermétalliques. Cependant la boule de cuivre a une dureté plus importante que celle de l'or, et donc les forces à mettre en jeu pour réaliser le soudage sont plus élevées ce qui peut entraîner des défaillances dans les interconnexions. D'autre part, le cuivre s'oxyde facilement à l'air, et requiert donc des conditions de mise en œuvre différentes.

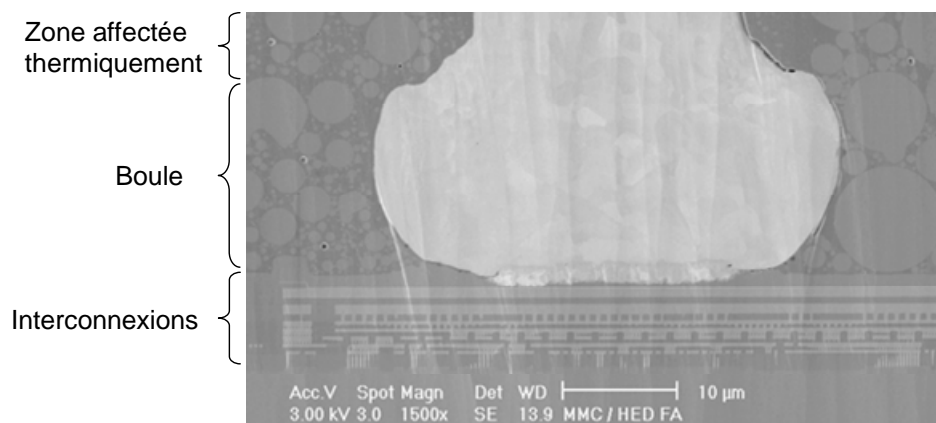


Figure 5. Observation au microscope électronique à balayage d'un assemblage filaire.

I.B.2. Méthodes de caractérisation

Afin de vérifier la fiabilité des connexions filaires, il existe plusieurs tests destructifs. Les deux techniques les plus couramment usitées sont le test de traction du fil (en anglais : « wire pull test ») et le test de cisaillement de la boule (en anglais : « shear test »).

I.B.2.a) Test de cisaillement de la boule

Ce test est la meilleure façon d'évaluer la qualité de l'interface entre la boule et le pad. Au cours de ce test, la boule est cisailée à une hauteur donnée, suivant la représentation schématique de la Figure 6. La force appliquée est mesurée, et suite à cela, la boule est soit cisailée, soit enlevée du pad. Dans ce dernier cas, elle peut également mener à l'arrachement d'une partie des interconnexions.

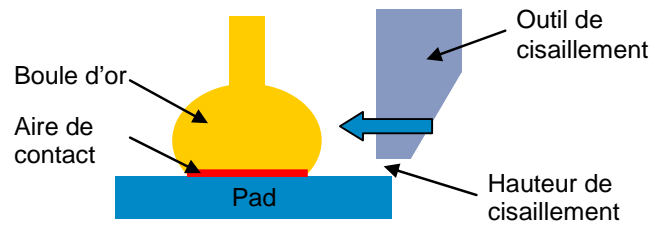


Figure 6. Représentation schématique du test de cisaillement de la boule.

I.B.2.b) Test de traction du fil

Le test de traction est réalisé pour contrôler l'opération de soudage. Les variables géométriques au cours de ce test sont décrites par la Figure 7. Ce test consiste à placer un crochet entre le fil et le pad, puis à appliquer un mouvement de traction verticale pour obtenir une force ainsi qu'un mode de rupture. Les modes de défaillances les plus courants sont la rupture dans la zone du fil affectée thermiquement, le détachement de la boule, caractérisant une faible liaison, et l'arrachement dans le pad. Dans le cas du premier mode de défaillance, le fil rompt juste au dessus de la boule dans la zone affectée thermiquement (recristallisation de la zone lors de l'application d'un arc électrique servant à former la boule au cours du processus de connexion filaire). L'arrachement du pad intervient dans les interconnexions, et constitue un mode de défaillance critique puisqu'il caractérise une architecture peu robuste.

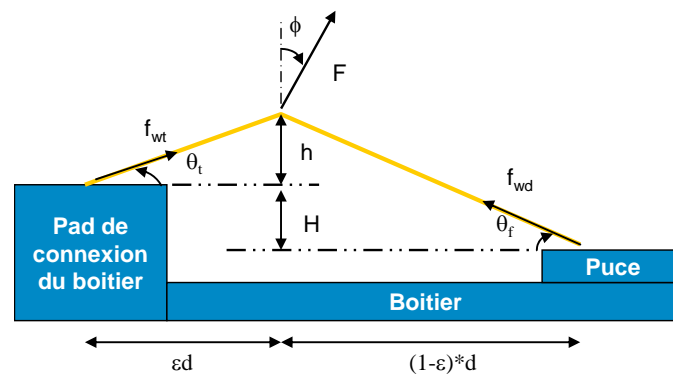


Figure 7. Variables géométriques du test de traction du fil, d'après [Harman, 1978].

I.C. Problèmes mécaniques induits

Au cours du tri électrique des puces et de l'assemblage filaire, diverses composantes de forces (compression, cisaillement, traction) sont mises en jeu, et peuvent provoquer l'apparition de fissures dans les couches d'oxyde situées au niveau des interconnexions. Ces fissures entraînent l'apparition de courts circuits dans les puces qui deviennent alors non-fonctionnelles. Les objectifs de cette thèse sont donc d'identifier les paramètres de test, d'assemblage entraînant l'apparition de ces défaillances, mais également de mettre en œuvre des moyens d'études rapides et fiables pour étudier diverses architectures de pad. Par ailleurs, l'étude et les résultats obtenus peuvent s'étendre à toutes les opérations effectuées au niveau des pads, que sont l'assemblage puce retournée ou « flip chip », l'assemblage via des colonnes de cuivre ou « copper pillar » ou encore l'assemblage grâce à des via traversant le silicium ou « through silicon via ».

II. Présentation des méthodes expérimentales et numériques mises en œuvre

Dans cette partie, les techniques d'imagerie utilisées pour caractériser les pads sont décrites. Elles permettent d'observer les fissures présentes dans les interconnexions, mais également de mieux comprendre les origines de ces défaillances. Ensuite, les méthodes expérimentales et numériques mises en place afin de reproduire les tests et assemblages effectués sur le pad sont introduites. Il s'agit d'une part de la nanoindentation, qui consiste en appliquer un chargement compressif sur la surface du pad. D'autre part, le modèle numérique par éléments finis est créé puis calibré grâce à des expériences dédiées, afin d'avoir un bon degré de confiance en les résultats obtenus.

II.A. Observation et reconstruction du pad en 3 dimensions

II.A.1.a) Observation de la surface du pad : Profilométrie optique

La profilométrie optique est une méthode d'observation basée sur le principe des interférences lumineuses. Une lumière blanche est envoyée sur l'échantillon à analyser, afin d'obtenir des franges d'interférences. Ces séries de frange d'interférences permettent, à la suite d'un balayage vertical de l'échantillon d'en observer sa surface en 3 dimensions. Un exemple de vue de l'aluminium après EWS est montré sur la Figure 3.

II.A.1.b) Reconstruction du volume du pad : FIB/MEB

La sonde ionique focalisée ou « Focussed Ion Beam » (FIB) est employée pour réaliser des coupes. Ces coupes peuvent être réalisées à des intervalles réguliers très faibles (56nm dans le cas présent). Le Microscope Electronique à Balayage (MEB) permet l'observation d'une surface ou une tranche grâce à un balayage d'électrons. Le couplage simultané des 2 techniques permet de cartographier un volume, comme cela est décrit en Figure 8. L'utilisation d'un logiciel permet ensuite de recouper les vues pour finalement obtenir une reconstruction suivant toutes les directions.

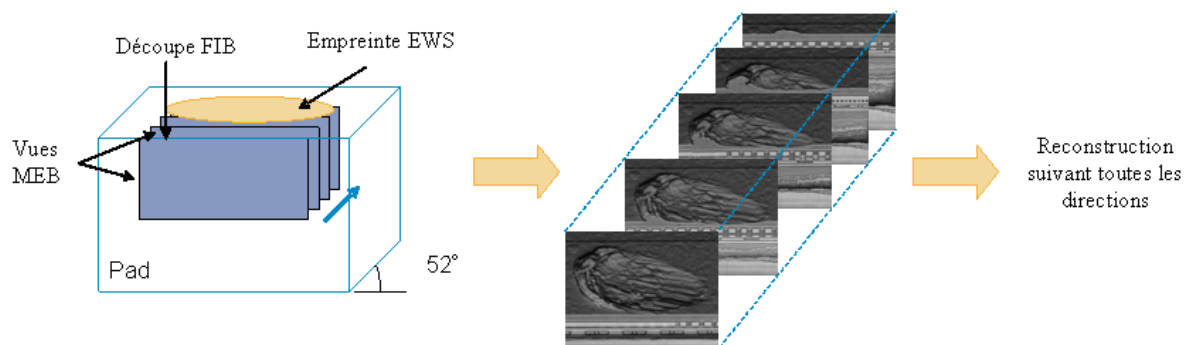


Figure 8. Principe de reconstruction FIB/MEB.

II.B. Evaluation de la robustesse des pads : nanoindentation

II.B.1.a) Principe et pointes utilisées

La nanoindentation est une technique récente permettant l'évaluation des propriétés mécaniques des couches minces [Doerner, 1986 ; Oliver, 1992]. Cette technique consiste à appliquer un chargement (souvent noté P) ou un déplacement (h) normal compressif, et semble donc adapté à la reproduction du test EWS particulièrement. Pour réaliser ces tests, un nanoindenteur MTS XP est utilisé. Au cours des expériences décrites, le chargement est

appliqué à travers 2 différentes géométries de pointes en diamant ayant 3 faces : la pointe Berkovich et la pointe cube corner (représentant le coin d'un cube). Les angles au sommet de ces pointes (θ) sont respectivement de 65.3° et 35.3° . Ces pointes sont décrites en Figure 9 et Figure 10.

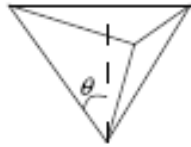


Figure 9. Pointe Berkovich.

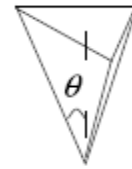


Figure 10. Pointe Cube corner.

II.B.1.b) Description du test

Un exemple de test d'indentation avec une pointe Berkovich est montré en Figure 11. La courbe se compose de 3 différentes parties. La première est une étape de chargement jusqu'à une force maximale P_{\max} ou un déplacement maximal h_{\max} . Elle est suivie d'une étape à chargement constant où l'enfoncement de la pointe est uniquement lié au fluage du matériau. Enfin, il se termine par une étape de déchargement afin de retirer l'indenteur. Au cours de cette étape, l'échantillon retrouve une partie de sa forme originale, dépendant de son élasticité. Le module d'Young et la dureté du matériau peuvent être extraits à la suite de ces tests suivant la méthode d'Oliver et Pharr [Oliver, 1992].

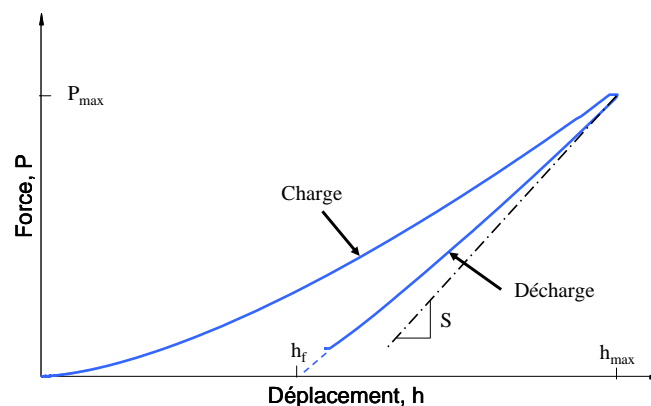


Figure 11. Courbe force-déplacement provenant d'une indentation avec une pointe Berkovich dans de la silice fondue.

Le mode CSM (« continuous stiffness mode ») est un mode permettant d'évaluer la raideur en continu, grâce à des cycles charge-décharge de très faible amplitude, afin d'obtenir le module élastique et la dureté du matériau durant la pénétration de la pointe.

II.C. Modèle numérique reproduisant l' EWS

II.C.1.a) Description du modèle éléments finis

De la même manière que les tests EWS, le modèle numérique se compose de 2 parties que sont la pointe et le pad, montrés en Figure 12. En raison du plan de symétrie, seulement la moitié de la pointe est simulée sous Ansys LSDyna. Les différents composants sont maillés avec des éléments quadrangle SOLID164 ayant un point d'intégration de Gauss. Les différentes couches composant le pad sont représentées par la Figure 13 a), et la géométrie du modèle associé est montrée en b). Les deux principales différences entre le pad réel et son

modèle numérique sont l'absence des couches barrières TaN/Ta (qui sera ultérieurement justifiée) et la non-représentation des détails dans le pad. Cela se caractérise par l'utilisation de la méthode d'homogénéisation des couches qui permet d'obtenir les propriétés équivalentes, suivant la méthodologie mise en place par Fiori et al. [Fiori, 2005]. Pour ces couches, des propriétés élastiques orthotropes sont donc introduites dans le modèle.

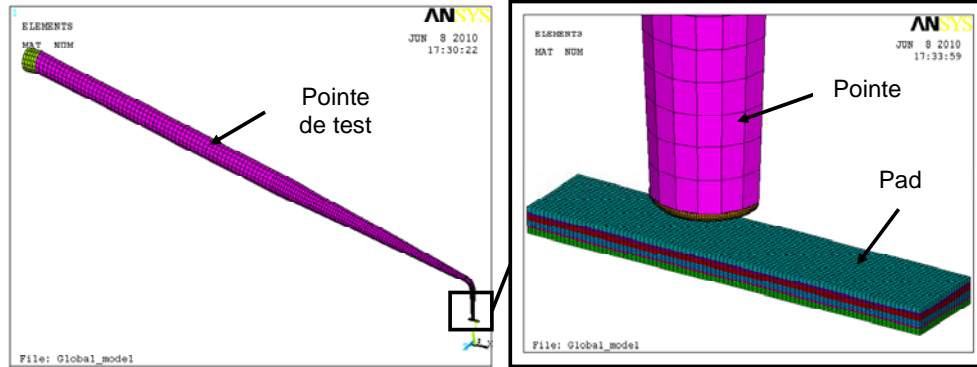


Figure 12. Modèle géométrique de la pointe de test (a) et zoom à l'échelle du pad (b).

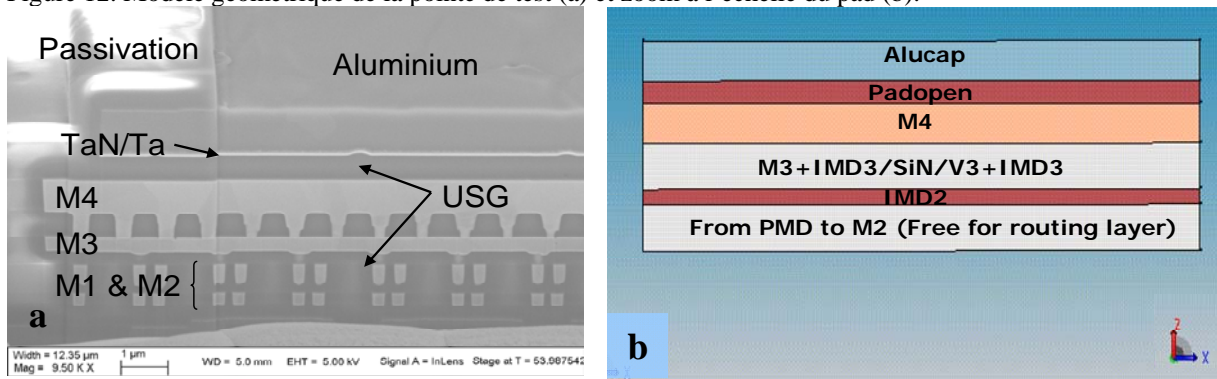


Figure 13. Vue MEB en coupe d'un pad CMOS 120nm (a) et vue du modèle géométrique correspondant (b).

II.C.1.b) Calibration du modèle

Des mesures des forces de réaction verticale et latérale sont effectuées au cours du probing par Technoprobe, le fournisseur de la carte à pointe décrite. Ces mesures sont effectuées sur l'appareil visible en Figure 14. La calibration du modèle a été initiée par Cacho et al. [Cacho, 2007]. Ainsi, la calibration du modèle est obtenue grâce à un bon accord entre les forces de réaction verticale et latérale mesurées et simulées, comme le montre la Figure 15. La courbe donnant la force de réaction verticale se compose tout d'abord d'une étape de chargement jusqu'à atteindre l'overdrive appliqué. Puis, un plateau permet d'effectuer la mesure électrique, et enfin la décharge jusqu'à la perte du contact entre la pointe et le plot de connexion.

. Ce résultat est valable pour les différents déplacements verticaux étudiés. Parallèlement à cela, les influences des paramètres numériques tels que le maillage sont vérifiés et il est montré que le modèle est capable de reproduire fidèlement les conditions de chargement du test réel.

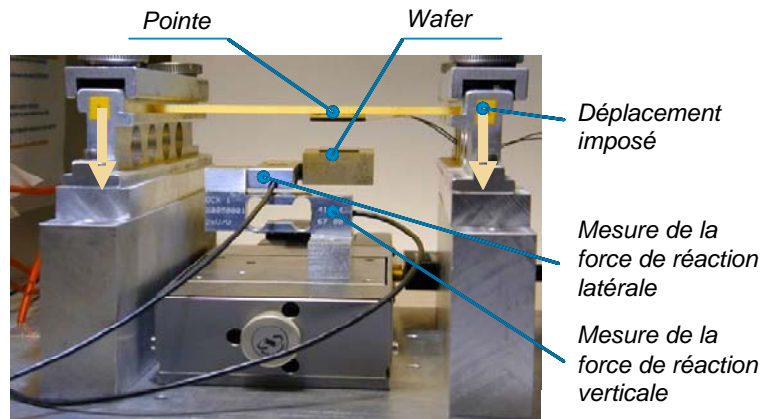


Figure 14. Appareil de mesure mis en place par Technoprobe pour obtenir les forces de réaction verticales et latérales.

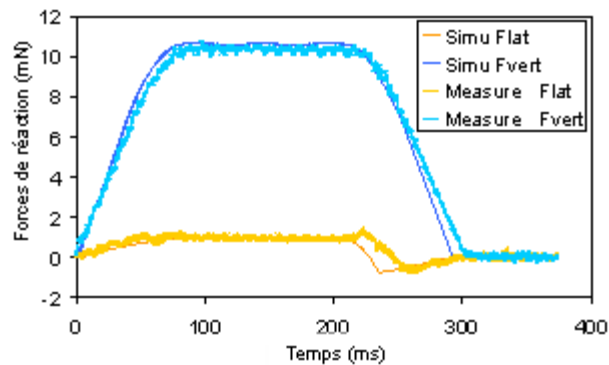


Figure 15. Forces de réaction mesurée et simulées pour un déplacement vertical de 50 μ m.

III. Effet des paramètres de test et d'assemblage sur l'apparition des défaillances

III.A. Paramètres critiques lié au test EWS

III.A.1.a) Résultats expérimentaux

Les paramètres décrits dans le paragraphe I.A.1 sont maintenant étudiés. Après avoir montré grâce à des simulations que les effets dynamiques (ex : la vitesse du plateau automatisé) sont négligeables dans le cas considéré, les 2 paramètres clés liés au tri des plaques sont évalués, à savoir le déplacement vertical (50 μ m et 90 μ m) et le nombre de passages (5 et 10 passages). Pour ce faire, des structures CMOS 120mn sont testées. L'endommagement de l'aluminium, c'est-à-dire la taille de la marque augmente fortement lorsque le déplacement vertical augmente, passant de 27 μ m à 40 μ m (Figure 16). L'effet du nombre de passage sur la longueur de la marque de test est négligeable. Cependant, il est montré que la largeur de la marque augmente en passant de 5 à 10 passages.

Afin de pouvoir dénombrer les fissures, le test d'intégrité est réalisé sur 2 ou 4 puces (148 pads par puce), afin d'avoir un échantillon représentatif. Les résultats sont montrés sur la Figure 17. L'augmentation de ces 2 paramètres conduit à une augmentation du pourcentage de pad présentant des fissures. Il est également démontré durant ces tests que les paramètres de test ont plus d'influence que ceux de design du pad, se référer également à [Roucou, 2010-A] pour plus de détails.

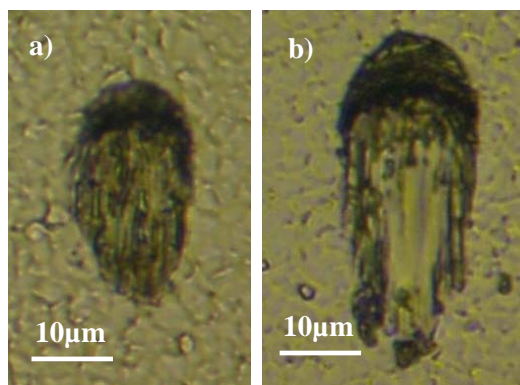


Figure 16. Vues de dessus par microscopie optique d'une marque suite au test avec un déplacement vertical de 50μm (a) et 90μm (b) pour 10 passages.

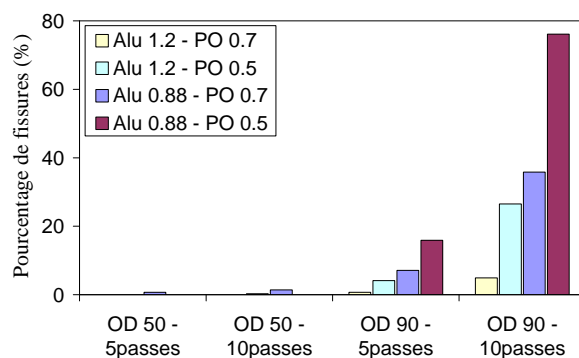


Figure 17. Pourcentage de fissures observées pour les différentes conditions de test (et différentes épaisseurs d'aluminium et d'oxyde pad open, voir également section IV.A.1.a).

III.A.1.b) Résultats numériques

Durant l'étape de calibration, l'augmentation du déplacement vertical de la plaque automatisée portant le wafer a été reliée à une augmentation de la force de réaction verticale. Durant les simulations numériques, il est montré que l'évolution des contraintes dans le pad peut être reliée à la force verticale appliquée, et de ce fait au déplacement vertical. De là, l'augmentation du déplacement vertical va induire une augmentation des contraintes compressives et Von Mises dans le pad [Roucou, 2010-A].

L'utilisation de propriétés élastiques pour l'aluminium en particulier limite l'étude au déplacement vertical puisque les contraintes dans les matériaux redeviennent nulles après un passage.

III.B. Connexion électrique par fil de cuivre

L'assemblage filaire est un mécanisme multi-étapes dont la représentation schématique des forces est montrée sur la Figure 18. Un film de cuivre enrobé de palladium est utilisé afin de limiter les problèmes liés à l'oxydation à l'air du fil, comme cela est démontré dans la publication de Vath et al. [Vath, 2009]. Le fil a un diamètre de 20μm, et ses propriétés mécaniques sont données dans l'article d'Uno et al. [Uno, 2009]. Les effets de la force de contact (notée également CF) ainsi que la puissance des ultrasons (notée également BP) sont évalués au cours d'un plan d'expérience qui comprend également des paramètres structuraux sur l'architecture des pads qui seront développés ultérieurement. Le premier paramètre varie entre 30mN et 50mN, tandis que le second varie entre 70mW et 90mW.

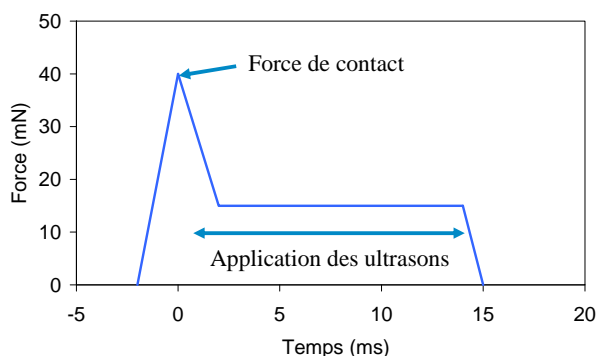


Figure 18. Représentation schématique des forces typiquement appliquées durant l'assemblage filaire.

Afin d'obtenir une bonne reproductibilité des résultats, les tests de traction sont effectués sur 3 puces contenant 15 fils et ceux de cisaillement sur 2 puces. Ainsi, 3240 fils sont tirés et 2160 boules sont cisailées.

III.B.1.a) Observation et mesure des boules

Le cuivre étant plus dur que l'or, au cours du soudage sur l'aluminium, une éclaboussure de ce dernier se produit. Ce phénomène n'est pas observé lorsque l'or est employé. La Figure 19 montre l'éclaboussure obtenue au cours des assemblages. Il est remarqué que la déformation plastique de l'aluminium se produit suivant la direction d'application des ultrasons. Cette déformation peut induire plusieurs problèmes mécaniques (dus à la réduction de l'épaisseur de l'aluminium sur le pad) et électriques (courts circuits dus à des ponts créés par les éclaboussures entre 2 pads voisins).

Au cours de cet assemblage filaire, les dimensions de la boule de cuivre (diamètre et hauteur) sont mesurées grâce à un microscope optique. L'incertitude sur les mesures ($0.4\mu\text{m}$) et le faible écart entre les valeurs obtenues font qu'il est difficile d'établir une corrélation entre les paramètres de soudage et la géométrie de boule, comme cela est montré sur la Figure 20 dans le cas du diamètre de la boule. Dans la littérature, pour certaines études, les variations observées sont également très faibles [Shah, 2010], tandis que dans d'autres, l'effet des ultrasons induit une augmentation du diamètre de la boule et une réduction de sa hauteur [Zhong, 2007 ; Xu, 2008].

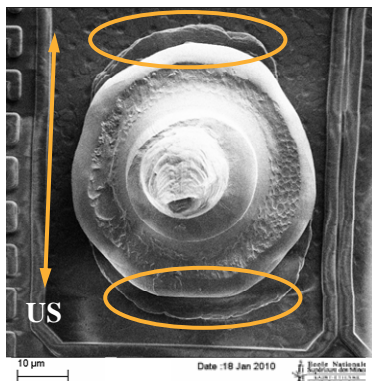


Figure 19. Vue MEB de dessus de la déformation plastique de l'aluminium au cours du processus.

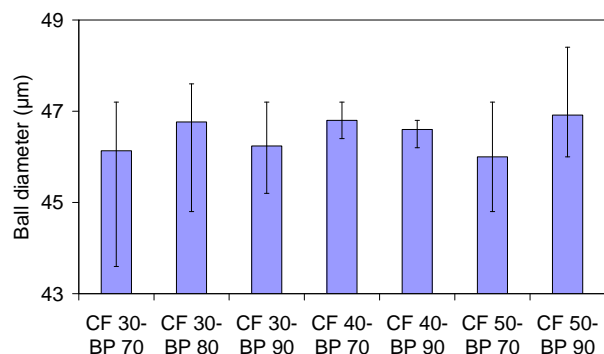


Figure 20. Effet des paramètres de soudage sur le diamètre de la boule de cuivre soudée.

III.B.1.b) Tests de traction et de cisaillement

Suite aux mesures des boules, un test de traction est effectué sur les fils de cuivre. Deux modes de défaillance sont obtenus : d'une part, les ruptures du fil au niveau de la zone affectée thermiquement (97% des fils), et d'autre part, un arrachement de certaines couches du pad (3% des fils, ou 98 cas sur 3240 fils). Ce dernier type de défaillance nous intéresse particulièrement puisqu'il permet de distinguer les structures en fonction de leur robustesse mécanique. De plus ces arrachements d'une partie des interconnexions correspondent à des résistances à la traction plus faibles que les ruptures dans la zone affectée thermiquement (de l'ordre de 78mN à 88mN). La répartition des cas d'arrachement de la structure en fonction des différentes conditions de soudage est montrée en Figure 21. Ainsi, il est montré que l'augmentation de la puissance des ultrasons et la diminution de la force de contact entraînent une augmentation du risque d'obtenir des arrachements. L'augmentation de la force de contact entraîne probablement une augmentation de l'aire de contact entre la boule et le pad. De là, lors de l'application des ultrasons, les forces sont distribuées sur une aire plus importante, et donc moins de contraintes sont subies par le pad. Cette hypothèse concernant

l'effet de la force de contact est également soulignée dans des travaux de Qin et al. [Qin, 2010] dans le cas de fils d'or. L'augmentation de la puissance des ultrasons peut être reliée à une augmentation de la contrainte de cisaillement dans le pad, ce qui est souligné par Shah et al. [Shah, 2009].

Parallèlement à ces tests, des tests de cisaillement de la boule sont effectués. Le mode de défaillance obtenu correspond à un détachement de la boule du pad. Ce mode de défaillance est dû, entre autre, à la faible formation des composés intermétalliques dans le cas de l'assemblage filaire avec du cuivre. Cette force est également reliée à une contrainte de cisaillement appliquée au pad, afin de limiter les effets de géométrie de la boule sur la valeur mesurée. Ainsi, dans le cas présent, des contraintes comprises entre 136MPa et 171MPa sont obtenues (voir [Harman, 1997] pour la méthode de calcul). Au cours de ces tests, l'augmentation de la force de contact induit une diminution de la force mesurée et de ce fait une diminution de la qualité du soudage. De ce fait, un équilibre doit être trouvé en ce qui concerne la force de contact à appliquer. Le soudage du cuivre est donc à différencier de celui de l'or pour lequel, l'augmentation de la qualité de la soudure est souhaitée et non corrélée à des effets d'arrachement de la structure.

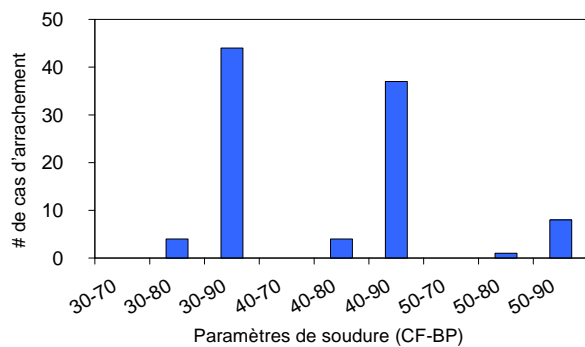


Figure 21. Nombre de cas d'arrachement de la structure observés pour les différentes conditions de soudage.

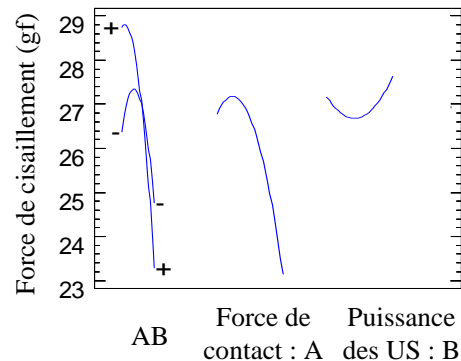


Figure 22. Effet des paramètres de soudage sur la force de cisaillement mesurée.

Dans le cas de la connexion électrique avec des fils de cuivre, les paramètres d'assemblage ont, comme dans le cas de l'EWS, plus d'influence sur les cas d'arrachement du pad que les paramètres de dessin de ces pads.

IV. Effet de l'architecture des pads sur robustesse mécanique

Dans le paragraphe précédent, l'effet des différents paramètres de test et d'assemblage a été souligné. Cependant, en raison de contraintes industrielles, leur modification n'est pas toujours possible, et la réduction des défaillances dans les architectures nécessite une optimisation de l'agencement de celles-ci. Dans ce paragraphe, les différentes modifications possibles au niveau de l'architecture des pads sont évoquées, de l'échelle la plus grande (épaisseur des couches), à la plus petite (design de la couche supérieure de cuivre).

IV.A. Effet de l'épaisseur des couches supérieures

La Figure 23 montre une représentation schématique de l'architecture étudiée. On considère durant cette étude les variations de l'épaisseur des deux couches supérieures du pad, à savoir l'aluminium et l'oxyde situé sous le pad (également appelé PO, pour « pad open oxide »). Les épaisseurs évaluées sont 0.88μm et 1.2μm pour l'aluminium et 0.5μm et 0.7μm pour l'oxyde.

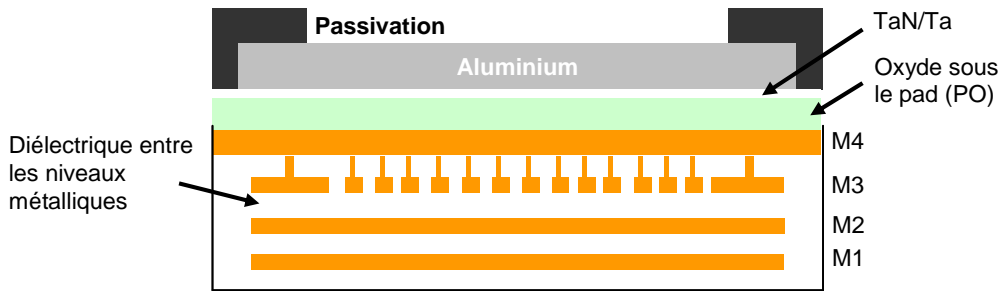


Figure 23. Représentation schématique des pads CMOS 120nm étudiés.

IV.A.1.a) Tests d'intégrité du pad et reconstruction en 3 dimensions

Les pourcentages de défaillances en fonction des différentes structures sont présentés sur la Figure 17. L'augmentation de l'épaisseur de la couche d'aluminium et de celle d'oxyde permet une diminution des risques de défaillances dans la couche d'oxyde située sous le pad. D'autre part, cet effet est cumulatif, puisque l'amélioration conjointe des 2 couches permet d'avoir la structure avec la meilleure robustesse mécanique. Ainsi, le classement suivant est obtenu. Il permet ultérieurement une comparaison avec les autres méthodes d'analyse.

$$\begin{matrix} \text{Alu} - B \\ \hline \text{PO} - B \end{matrix} < \begin{matrix} \text{Alu} - B \\ \hline \text{PO} - H \end{matrix} < \begin{matrix} \text{Alu} - H \\ \hline \text{PO} - B \end{matrix} < \begin{matrix} \text{Alu} - H \\ \hline \text{PO} - H \end{matrix} \quad (1)$$

Une reconstruction du pad suivant la méthode décrite dans le paragraphe II.A.1.b) est réalisée. Cela permet d'une part d'obtenir des vues MEB en coupe du pad (Figure 24), et d'autre part de reconstruire et donc visualiser les couches en vue de dessus (Figure 25). Une importante déformation plastique est observée pour l'aluminium, menant à la formation d'un amas en front de pointe de test. Des fissures importantes sont observées dans l'oxyde PO (reconstruction a)), dans une zone où l'aluminium du pad a été enlevé durant le test EWS. Des fissures de plus petite taille sont également observées au niveau des singularités géométriques, dans les couches d'oxyde situées plus profondément dans le pad. Enfin, la couche supérieure de cuivre (M4) est également déformée plastiquement à l'issue du test (reconstruction b)). De là, l'effet prouvant la ductilité de l'aluminium et du cuivre afin de réduire des contraintes mécaniques dans le pad est souligné.

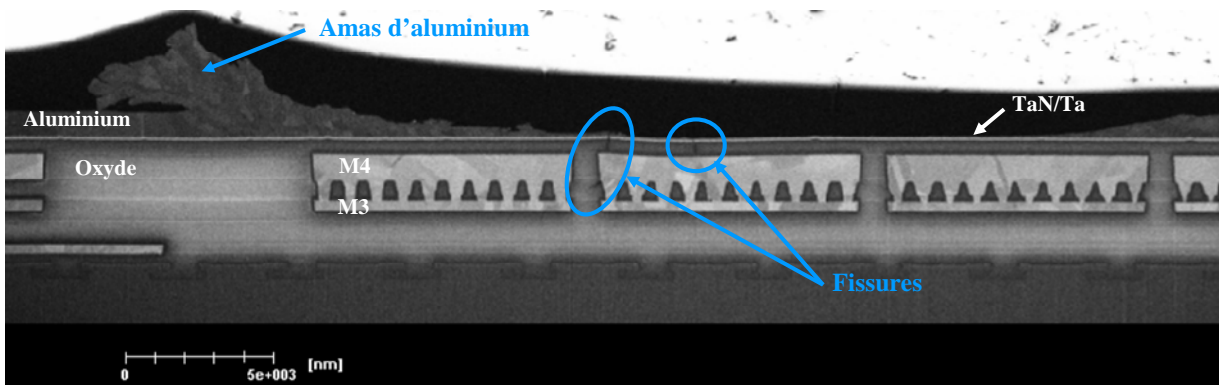


Figure 24. Vue MEB en coupe du pad.

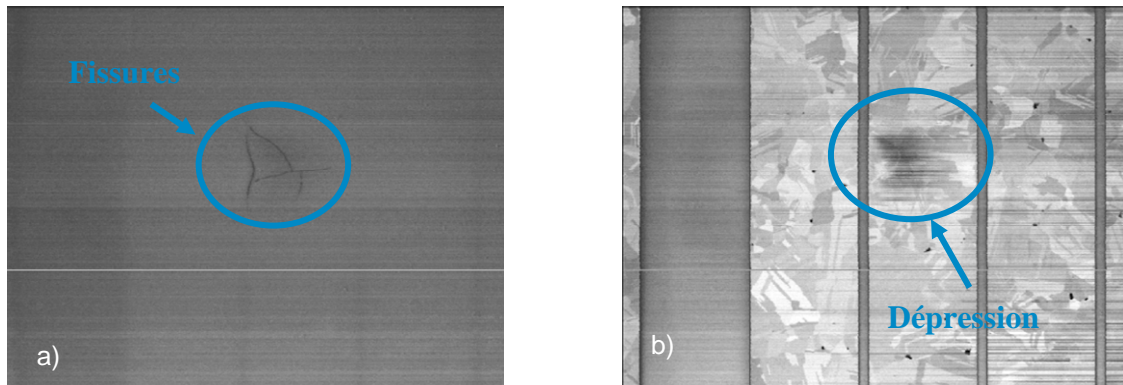


Figure 25. Images reconstruites suivant la technique FIB/MEB de la couche d'oxyde située sous le pad (a) et de la couche supérieure M4 du cuivre (b)

IV.A.1.b) Résistance à la compression : nanoindentation

L'application de la technique de nanoindentation aux architectures de pad est rendue difficile du fait de la présence d'un empilement multicouches. Néanmoins, des tests de nanoindentation sont effectués sur les architectures avec les pointes Berkovich et cube corner.

L'utilisation de la pointe Berkovich permet la reproduction de défaillances similaires à celle du test EWS, comme cela est montré par la Figure 26. Les mêmes observations en termes de plasticité de l'aluminium et du cuivre peuvent également être faites. La fissure dans l'oxyde de la couche sous le pad apparaît à l'endroit où celui-ci est le plus courbé, et la localisation de cette fissure est probablement dépendante de l'épaisseur de l'oxyde et de la pointe utilisée pour le test. Les résultats de la pointe Berkovich sont semblables à ceux de la pointe cube corner, mais légèrement moins dispersés, et ne sont donc pas développés.

Ensuite, 15 tests avec la pointe cube corner sont réalisés sur chaque structure afin de vérifier la reproductibilité du test, qui est obtenue comme cela est montré sur la Figure 28 a). La force appliquée est fixée à 15mN. Les déplacements atteints à la fin du chargement sont comparés pour chaque structure, considérant qu'ils représentent leur résistance à la compression. Cependant, la présence de l'aluminium fausse légèrement le classement des architectures de pad puisque la pointe pénètre très facilement au travers (voir les courbes sur la Figure 27). Sur ces courbes de nanoindentation avec la pointe cube corner, un léger décrochement de la courbe apparaît au niveau de l'oxyde. Ce petit plateau caractérise la présence d'une microfissure. De là, un nouvel indicateur est introduit : $h \cdot dh/dp$, qui dépend donc de l'inverse de la pente, et permet une meilleure observation de ces fissures pour en extraire ensuite le chargement critique correspondant et la profondeur d'apparition, comme cela est montré sur la Figure 28 b). L'utilisation de ce nouveau paramètre permet de contrebalancer l'effet de l'aluminium et d'obtenir un classement en fonction de la capacité des structures à résister à la compression. Ce classement est le même que celui obtenu dans le cas des tests réels (EWS), avec une inversion entre les 2 structures avec une faible et une forte épaisseur. D'autres informations concernant cette méthode d'analyse sont également disponibles dans une précédente publication [Roucou, 2010-B].

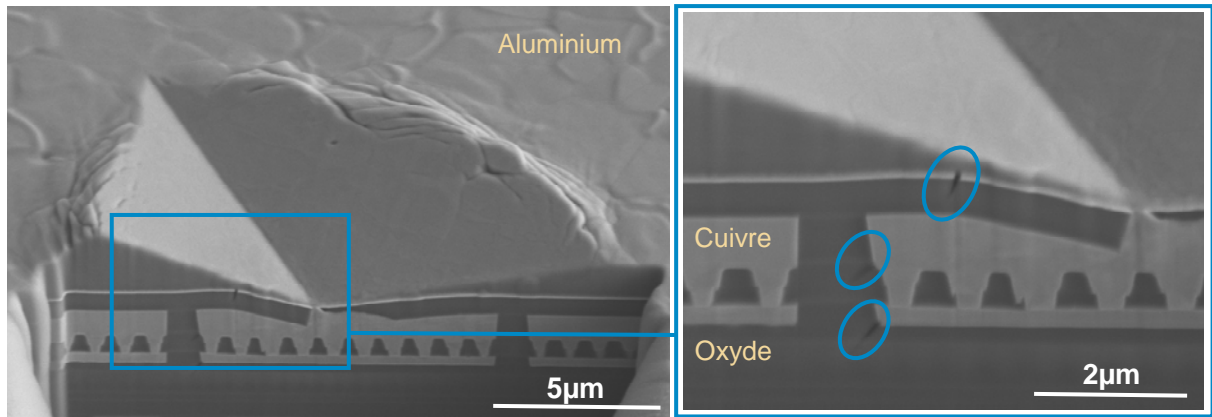


Figure 26. Vue MEB en coupe du pad après coupe FIB suivant une arrête de l'indenteur, et zoom sur la localisation des fissures.

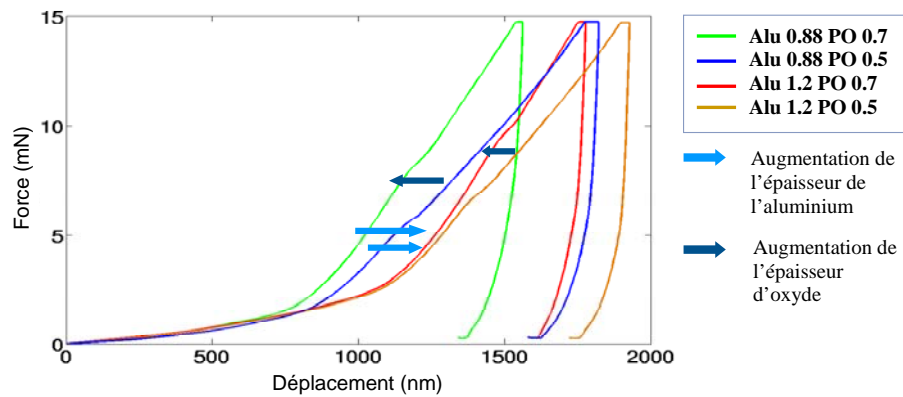


Figure 27. Courbes force-déplacement moyennées pour les différents empilements

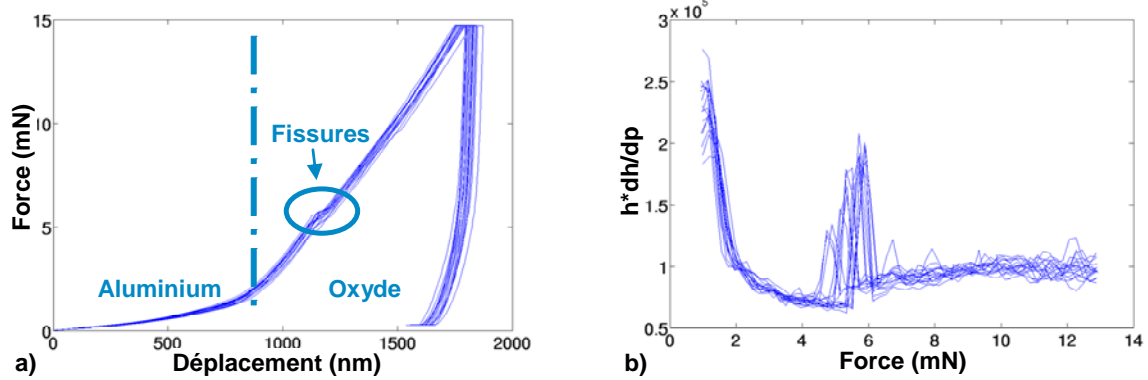


Figure 28. Courbe force déplacement pour une structure donnée (a), et courbe $h \cdot dh/dp$ en fonction de la force pour la structure correspondante (b).

IV.A.1.c) Analyse des contraintes dans le pad : modèle numérique

Il a été montré précédemment que la contrainte maximale dans le pad est obtenue durant le plateau à la fin de l'étape de chargement (cf. Figure 15). De là, les contraintes compressive et de Von Mises dans les différentes couches sont extraites durant une de ces sous-étapes de la simulation, et les résultats sont regroupés sur la Figure 30. Les contraintes dans les couches d'oxyde (PO et M3/IMD3) sont particulièrement étudiées puisque les fissures ont été localisées dans ces couches. L'augmentation de l'épaisseur des couches d'aluminium et d'oxyde PO implique une réduction des contraintes dans la structure, et de ce fait une réduction des risques d'apparition de fissures. Ainsi, un classement entre les différentes structures peut être établi sur la base des contraintes maximales simulées. Il est identique à celui des tests EWS (équation (1)).

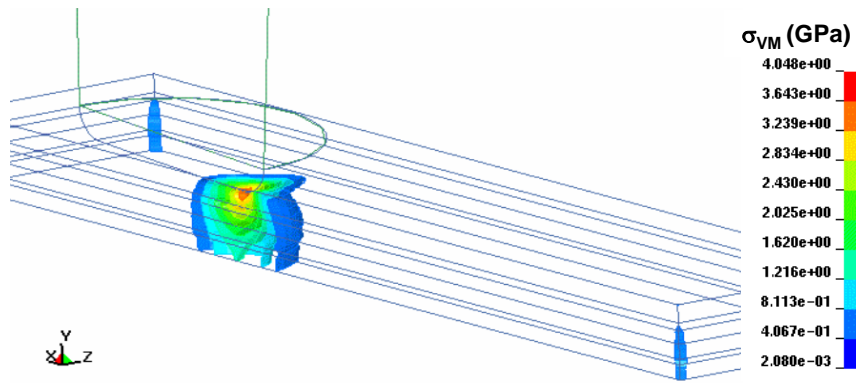


Figure 29. Contraintes de Von Mises dans le pad.

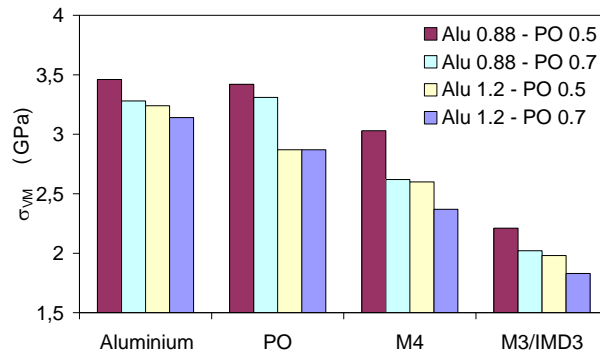


Figure 30. Contraintes Von Mises maximales obtenues dans les couches supérieures pour les différentes structures.

Ainsi, dans ce paragraphe, il a été montré que la nanoindentation et la simulation sont toutes deux capables de reproduire les résultats obtenus au cours de tests de tri des plaques, même si quelques limitations en ce qui concerne la nanoindentation ont été remarquées, et seront discutées ultérieurement.

IV.B. Effets de la densité de cuivre et de la présence des vias

Dans ce paragraphe, les effets de la densité de cuivre dans les couches inférieures de cuivre, ainsi que la présence de via entre les niveaux de cuivre et au niveau de l'oxyde situé sous le pad sont évalués au cours du plan d'expérience sur l'assemblage filaire avec du cuivre initié précédemment.

IV.B.1.a) Description des architectures étudiées

Les architectures étudiées comprennent 4 niveaux de cuivre (CMOS 120nm). La représentation schématique des différentes variétés d'architecture étudiées est décrite sur la Figure 31. Les différentes structures sont montrées dans la Tableau 1.

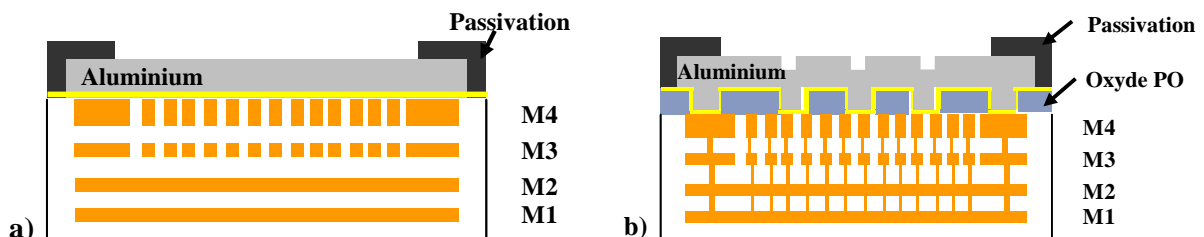


Figure 31. Représentation schématique de la structure sans via, ni via sous l'aluminium (a), et avec vias et vias dans l'oxyde sous le pad (b).

Structure de pad	Densité de métal (M1 à M3)	Présence de via	Présence de via dans l'oxyde sous le pad
A	30%	Non	Oui
B	80%	Oui	Oui
C	80%	Non	Non
D	80%	Non	Oui
E	30%	Oui	Non
F	30%	Non	Non
G	80%	Oui	Non
H	30%	Oui	Oui

Tableau 1. Plan d'expérience mis en place durant l'étude.

IV.B.1.b) Tests de traction et observation des cas d'arrachement

Afin d'obtenir des tendances fiables, pour chaque structure et chaque condition d'assemblage (voir paragraphe III.B), 45 fils sont tirés. Comme précédemment, l'objectif est d'obtenir des cas d'arrachement du pad afin de comparer la résistance mécanique des pads. Les résultats sont réunis dans le tableau ci-dessous. A partir de cela, un classement des diverses architectures peut être établi. Il est notamment montré que la présence de via sous le pad est particulièrement critique en ce qui concerne l'apparition des arrachements. A contrario, la densité des couches inférieures du pad, et la présence de via, n'est pas montré dans le cas présent comme étant un paramètre déterminant. Une des raisons pouvant expliquer cette dernière observation est liée à la présence d'une couche de cuivre de forte densité (80%) entre le pad et ces couches (c.-à-d. M4), et beaucoup plus épaisse, jouant un rôle d'écrantage.

CF/BP Pads	30 70	30 80	30 90	40 70	40 80	40 90	50 70	50 80	50 90	Total
A	0	2	10	0	1	6	0	0	0	19
B	0	1	7	0	2	5	0	0	3	18
C	0	0	2	0	0	1	0	0	0	3
D	0	1	15	0	0	12	0	0	1	29
E	0	0	0	0	0	0	0	0	0	0
F	0	0	1	0	0	0	0	0	0	1
G	0	0	0	0	0	1	0	0	0	1
H	0	0	9	0	1	12	0	1	4	27
Total	0	4	44	0	4	37	0	1	8	98

Tableau 2. Nombre de cas d'arrachement du pad parmi les 45 fils tirés pour chaque paramètre étudié.

Puis les arrachements sont observés au MEB et la composition des couches supérieures est analysée grâce à une analyse dispersive en énergie (communément appelé EDX pour « Energy dispersive X-ray spectrometry »). De là, il faut distinguer 2 cas, celui des pads contenant de l'oxyde et des vias sous le pad, des autres. Un exemple typique des cas d'arrachement pour le premier type de pad est montré sur la Figure 32. La delamination entre les couches s'est produite au dessus de M4, au niveau des interfaces Cu/SiO₂ ou Cu/Si₃N₄. De ce fait, il semble que la présence des nombreuses singularités géométriques induites par la présence des vias et de l'oxyde est la raison du fort nombre d'arrachements observés. Au vue des pads arrachés, le mécanisme d'arrachement semble se produire en deux étapes : tout d'abord, une delamination se produit dans le pad, au niveau des interfaces incriminées, puis se propage pour atteindre la surface, ce qui explique le léger surélèvement des couches en bord d'arrachement. Pour les structures ne présentant pas ces singularités géométriques, l'arrachement se situe plus profondément dans la structure, plus précisément juste au-dessus ou au-dessous de la couche M3, comme cela est observé sur la Figure 33. Dans ce cas, le

même type de mécanisme est suspecté, mais s’initiant cette fois à l’interface entre le cuivre du niveau M3 et une des couches à proximité.

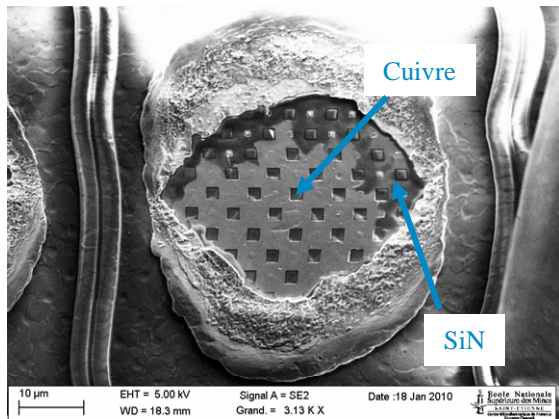


Figure 32. Vue MEB de dessus d’un pad de type D : l’arrachement est observé au dessus de M4.

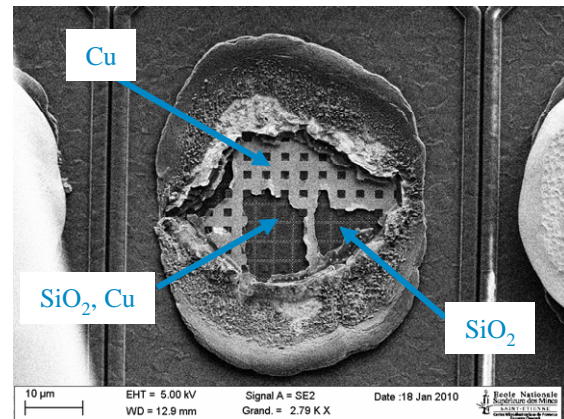


Figure 33. Vue MEB de dessus d’un pad de type C, l’arrachement est situé au-dessus et au-dessous de M3.

IV.B.1.c) Tests de nanoindentation

Parallèlement aux tests de traction, des tests de nanoindentation sont effectués sur les architectures. Afin d’avoir une relativement bonne reproductibilité des résultats, 2 séries de 15 tests sont effectués sur chaque structure, et les enfoncements maxima atteints à la fin du chargement pour les différentes structures sont montrés sur la Figure 34. De là, il est clairement montré que les structures ayant des vias dans l’oxyde sous l’aluminium résistent mieux à la compression. Cela est dû à la différence entre les propriétés matériaux (E, H) du cuivre et de l’oxyde. En effet, dans un cas, l’indentation évalue un empilement aluminium sur oxyde, tandis que dans l’autre cas, il s’agit d’un empilement aluminium sur cuivre.

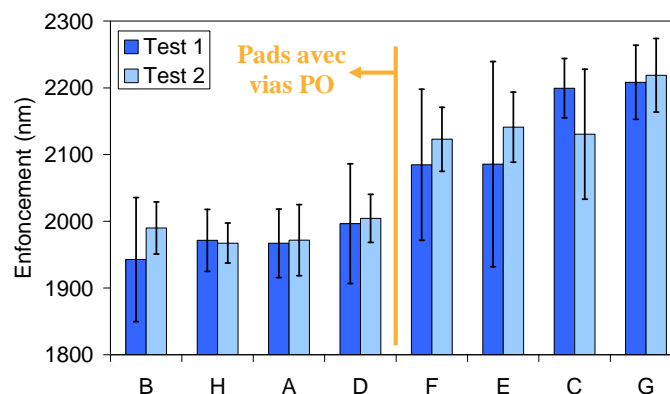


Figure 34. Déplacements maxima atteints après le chargement pour les diverses structures étudiées.

Ainsi, dans certains cas, en particulier lorsque différentes couches sont évaluées, il est difficile par des tests de nanoindentation de reproduire les résultats d’assemblage filaire. Néanmoins, il est montré en annexe de la thèse que des résultats coïncidents avec les 2 techniques peuvent être obtenus. La nanoindentation constitue un chargement simplifié par rapport à l’assemblage filaire, puisqu’uniquement de la compression est appliqué au niveau du pad, tandis que les résultats des tests de traction mettent en jeu des contraintes de compression et de cisaillement durant le processus, puis des contraintes de traction (principalement) durant le test.

IV.B.1.d) Etude sur des architectures faibles permittivités

Le besoin de miniaturisation dans l'industrie des semi-conducteurs s'accompagne de l'utilisation de nouveaux matériaux à faible permittivité (couramment appelés low-k en anglais), qui sont fragiles, et possèdent des propriétés matériaux très faibles [Moore, 2001 ; Wu, 2006]. A partir de ce constat, les paramètres étudiés précédemment le sont également grâce à la nanoindentation pour une technologie CMOS 45nm. Les tests faits avec la pointes Berkovich montrent d'une part que la présence de via dans l'oxyde sous le pad implique une réduction du déplacement maximum atteint à la fin du chargement. Dans le cas présent, le cas ne présentant pas de via comporte de l'oxyde entre l'aluminium et le cuivre. Il est donc montré que la présence de via sous le pad réduit la résistance de celui-ci à la compression.

D'autre part, il est montré qu'augmenter la densité du cuivre entre les couches comprenant des oxydes basse permittivité (situé dans les premiers niveaux, proche du silicium) permet d'augmenter la résistance de la structure face à un chargement compressif. La différence entre les propriétés matériaux du cuivre et les matériaux basse permittivité explique cette inversion de tendance entre le cuivre et l'oxyde.

IV.C. Effet de l'agencement de la couche de cuivre

IV.C.1.a) Etude sur des architectures faibles permittivités

L'effet de 2 types d'agencement de la couche supérieure du cuivre sur la robustesse du pad est étudié : des lignes de cuivre et des lignes larges dans lesquels des trous d'oxyde sont disposés à intervalle régulier (également appelés « slots »). Ces agencements, ainsi que les fissures typiques retrouvées à l'intérieur sont observés sur la Figure 35. Les pourcentages de défaillance dans les différentes structures sont montrés sur la Figure 36. Ainsi, il est montré que pour réduire le risque d'observer des fissures, il faut privilégier les structures avec des slots, de petite taille si possible, aux structures avec des lignes. Les effets des contraintes résiduelles et de la microstructure du cuivre sont suspectés comme étant à l'origine de ces différences. L'effet de la localisation des plateaux contenant les slots est également souligné, et il est préférable d'éloigner la ligne d'oxyde entre les plateaux de la zone de l'EWS. Ainsi le classement suivant est obtenu :

$$\text{Lignes} < 2\mu\text{m Slots} < 1\mu\text{m Slots} < 1\mu\text{m Slots with plateaux de cuivre décalés} \quad (2)$$

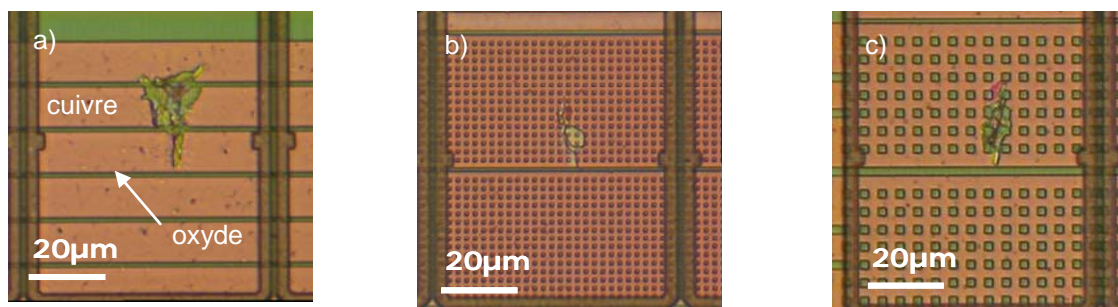


Figure 35. Vues de dessus, au microscope optique, des différents pads montrant des fissures dans l'oxyde : pour une structure avec des lignes (a), avec des trous d'oxyde de 1µm dans des plateaux de cuivre (b), et des trous de 2µm (c).

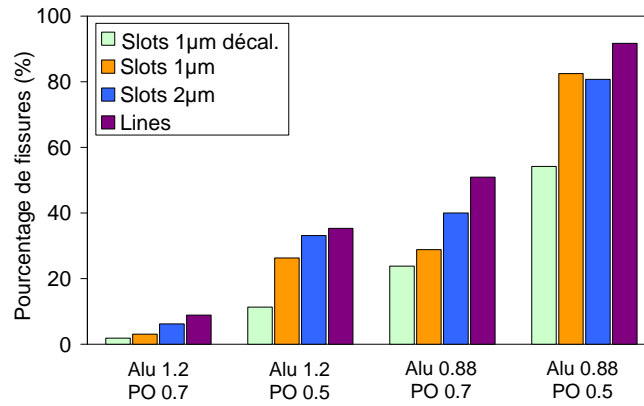


Figure 36. Pourcentage de fissures observées pour les différentes architectures.

IV.C.1.b) Nanoindentation

Parallèlement, de la nanoindentation avec la pointe cube corner est réalisée sur ces plots de connexion. A partir des indicateurs précédemment utilisés, à savoir la profondeur atteinte à la fin du chargement, le déplacement et la force caractérisant la fissure, un classement à partir des données de la nanoindentation peut être établi, et un classement similaire à celui des tests réels est obtenu hormis l'effet du décalage du plateau. L'analyse locale induite par la pointe cube corner est à l'origine de cette différence, ainsi que les effets de cisaillement induits par les tests réels.

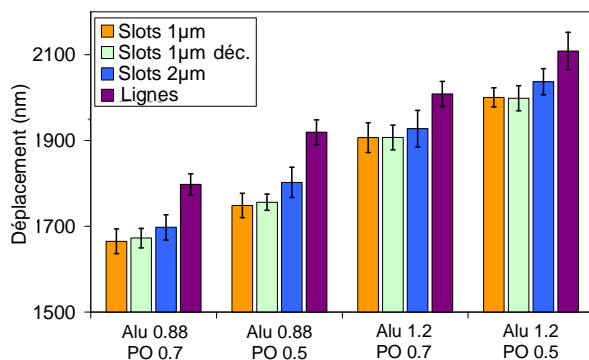


Figure 37. Profondeur maximale atteinte pour les différents échantillons.

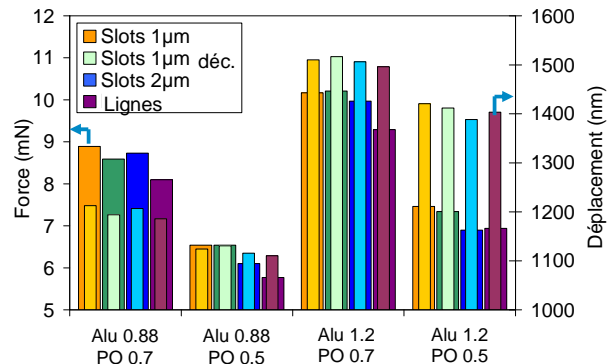


Figure 38. Force et déplacement relatifs aux fissures dans l'oxyde.

IV.C.1.c) Etude numérique

Le modèle éléments finis est utilisé pour évaluer les contraintes dans les couches de la structure. Néanmoins, l'utilisation de propriétés homogénéisées pour le cuivre ne permet pas de distinguer les structures contenant des slots entre elle, puisqu'elles ont les mêmes densités de métal et répartition dans l'espace, ce qui implique des propriétés orthotropiques équivalentes. Il est donc uniquement possible de comparer les lignes de cuivre aux slots. Cette dernière structure est reliée à des contraintes plus faibles dans le pad et donc un plus faible risque d'observer des fissures.

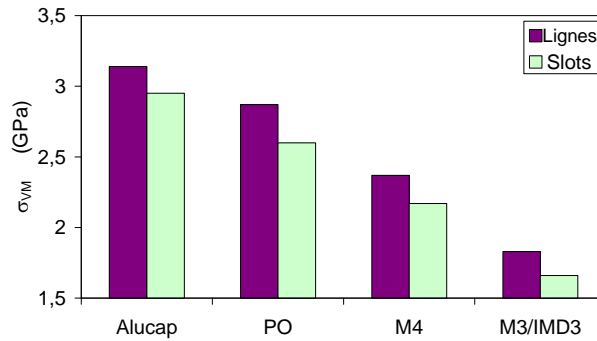


Figure 39. Contraintes Von Mises maximales dans les différentes couches dans le cas de l'empilement le plus épais.

Finalement, dans cette partie, la capacité de la nanoindentation et du modèle éléments finis à reproduire les résultats des tests réels effectués sur les pads a été démontrée. Ces 2 méthodes constituent donc des alternatives moins coûteuse et plus rapide pouvant permettre une amélioration des architectures de pad. Des règles de dessin concernant les empilements et les agencements des différentes couches constituant le pad ont également été obtenues.

V. Adaptation au contexte industriel

Les travaux se déroulant dans un contexte industriel, il est nécessaire d'avoir des modèles rapides pouvant donner des indications fiables lors du développement technologique ou pour la résolution de crises. De ce fait, des règles de tests doivent être définies dans le cadre de la nanoindentation, et un nouveau modèle éléments finis, plus rapide et dédié à l'évaluation des architectures de pad doit être développé au niveau numérique.

V.A. Méthodologie pour l'étude des architectures de pad

V.A.1.a) *Comparaison entre les pointes Berkovich et cube corner*

L'empreinte laissée par la pointe Berkovich est beaucoup plus grande que celle laissée par une pointe cube corner du fait de la géométrie des 2 pointes. Pour atteindre un enfoncement donné, la force à appliquer sur une pointe Berkovich est également plus importante (de l'ordre de 10 fois), comme cela est montré sur la Figure 40. De là, localement, la pointe cube corner va induire des contraintes et des déformations beaucoup plus importantes, comme cela a été démontré par Bucaille et al. [Bucaille, 2003]. Néanmoins le volume déformé plastiquement sera plus important dans le cas d'une pointe Berkovich. Ainsi, le choix de la pointe influe sur le volume indenté et donc fortement sur la réponse obtenue. La Figure 41 souligne cet effet puisque l'interface entre l'aluminium et l'oxyde est clairement marqué lors d'une indentation avec une pointe cube corner. Enfin, le retour élastique suite est plus important pour une pointe Berkovich ($h_f/h_{max}=0.86$) que pour une pointe cube corner ($h_f/h_{max}=0.93$).

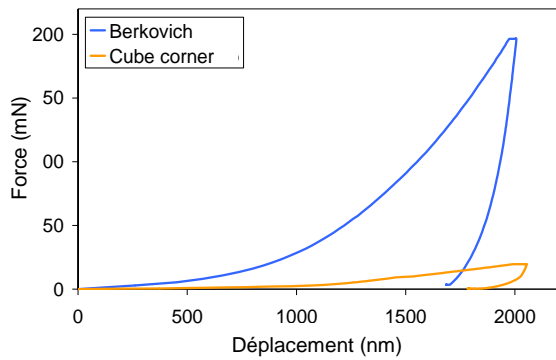


Figure 40. Courbe donnant la force en fonction du déplacement pour les 2 types de pointe sur des échantillons de CMOS 45nm.

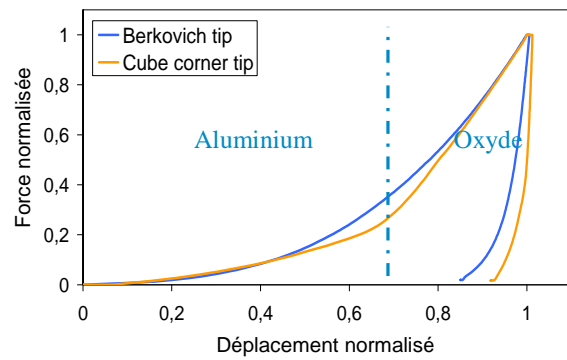


Figure 41. Courbes force-déplacement précédentes normalisées

V.A.1.b) Règles d'utilisation pour l'étude des pads

Les différences entre les 2 pointes ont été soulignées dans les divers paragraphes précédents. Dans quelques cas, des réponses légèrement différentes ont été obtenues avec les 2 pointes. De ce fait, il apparaît nécessaire de clarifier le choix de la pointe pour l'étude des pads. D'une manière générale, il est préférable de coupler les 2 études, cependant, lorsqu'un choix de pointe est nécessaire, il faut tenir compte de l'architecture de l'échantillon, et les règles décrites dans le diagramme suivant sont conseillées.

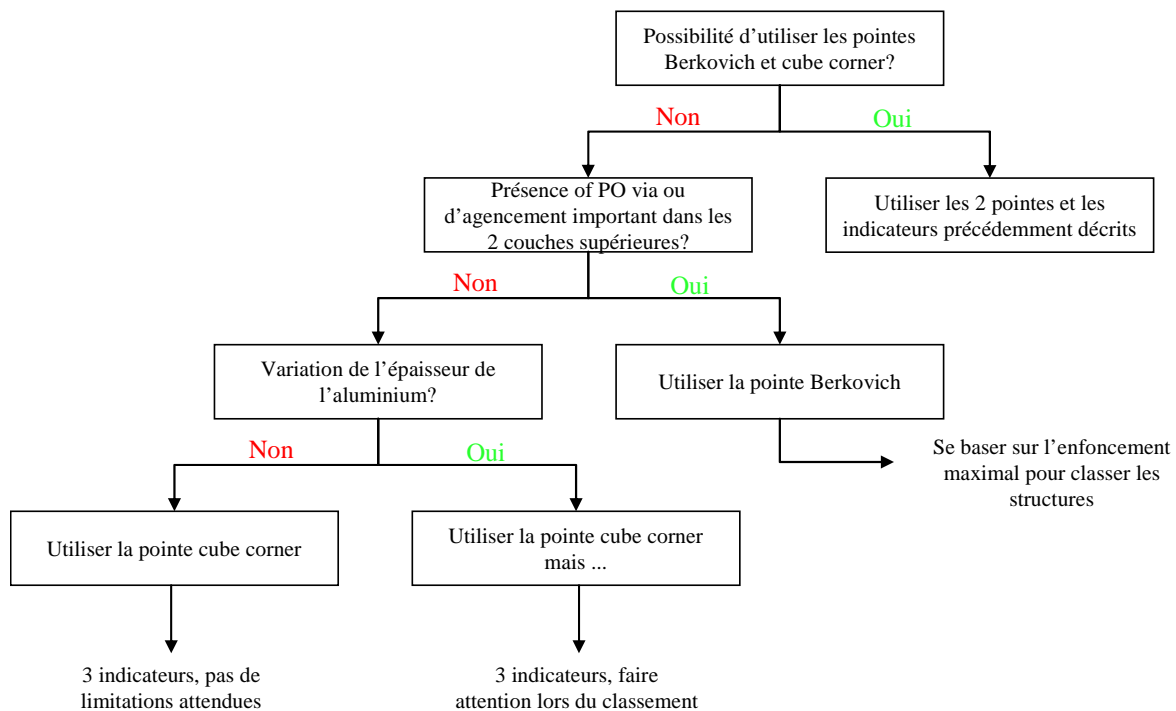


Figure 42. Diagramme indiquant le choix du type de pointe en fonction de l'architecture du pad étudiée.

V.B. Développement d'un nouveau modèle numérique reproduisant l'EWS

V.B.1.a) Objectifs du modèle

Il a été montré que le modèle explicite développé précédemment pour reproduire le tri électrique des wafers est capable de donner la plupart des tendances obtenues, qu'il s'agisse de paramètres liés au test ou à l'architecture. Cependant, ce modèle requiert un fort temps de calcul (environ 100 heures) et montre certaines limitations en ce qui concerne l'étude des designs de pad. De là, il est nécessaire de développer un nouveau modèle qui tient compte de tous les résultats expérimentaux et numériques démontrés. Ce modèle est dédié à l'étude des pads et doit par conséquent permettre d'introduire des détails très petits tels que les vias.

D'autre part, afin de réduire le temps de calcul, le modèle élément fini avec un schéma de résolution implicite est requis. De même, uniquement le pad est représenté, ce qui permet de s'affranchir des problèmes de convergences liés au contact. L'utilisation d'un modèle en 3 dimensions est justifiée par la complexité des designs du pad.

V.B.1.b) Développement du modèle 3D

Le développement du modèle en 3 dimensions, nécessite au préalable une étude en 2 dimensions afin de définir les conditions de chargement. Le modèle est développé sous le logiciel ANSYS multiphysics [Ansys, rel. 11]. Ainsi, pour reproduire le chargement, un contact de Hertz est utilisé. L'utilisation de ce contact est justifiée par la déflexion de la pointe observée dans le modèle éléments finis explicite, accompagné d'une réduction de l'aire de contact au cours du chargement lors de l'EWS. De même des abaques, basées sur les tests expérimentaux servant à la calibration du modèle explicite sont utilisées afin d'obtenir les forces de réaction verticale et latérale pour les différents paramètres de test. Après application, il est montré que les conditions de chargement multi-étapes employées aboutissent à des maxima de contraintes dans l'oxyde sous le pad, conformément aux attentes. Le modèle symétrique maillé en 3 dimensions est montré sur la Figure 43.

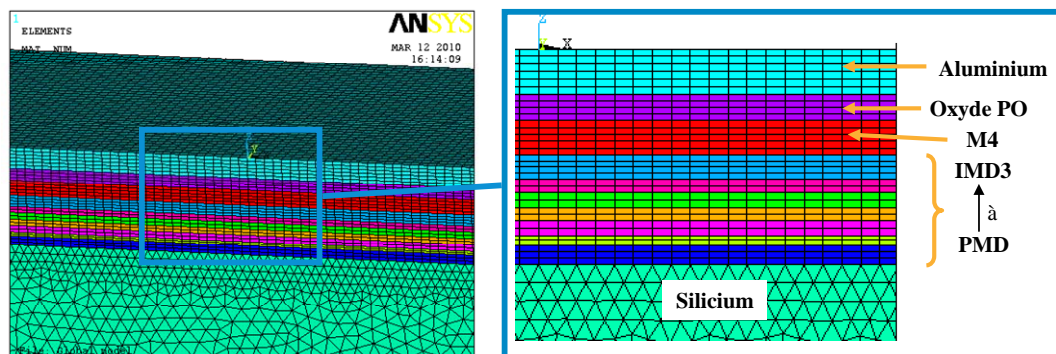


Figure 43. Modèle éléments finis du pad en 3 dimensions

Afin de reproduire les bonnes conditions de chargement, différents profils sont étudiés. Ainsi l'objectif est dans un premier temps de reproduire la marque de l'EWS, en conférant à l'aluminium des propriétés élasto-plastiques (limite d'élasticité : 200MPa et module tangent : 6GPa). La marque reproduite à une longueur plus faible que la marque réelle, ce qui est probablement dû au fait que dans le cas des tests réels, la zone finale ne représente pas réellement le chemin de la pointe sur le pad, c'est-à-dire qu'une partie de l'aluminium en surface du pad n'est pas déformée plastiquement lors du passage de la pointe. Ce résultat a également été observé par Strom et al. [Strom, 1998].

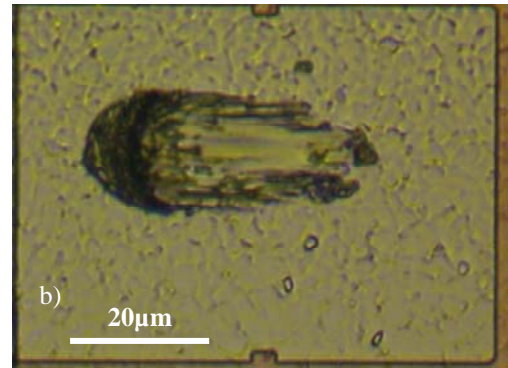
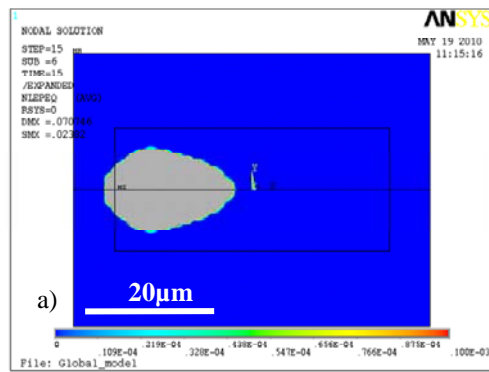


Figure 44. Déformation plastique cumulée dans l'aluminium (a) et vue de dessus d'une marque d'EWS (b).

Les paramètres de tri des plaques sont ensuite évalués en gardant les propriétés élasto-plastique pour l'aluminium. Il est montré que l'augmentation du déplacement vertical induit une augmentation de la zone déformée plastiquement, conformément aux mesures expérimentales. L'effet du nombre de passages est également étudié. Il est démontré que les variations dues au second passage pour le modèle défini sont faibles. Des contraintes résiduelles sont observées dans l'aluminium et l'oxyde à la fin du premier passage (voir Figure 45 b)). Les effets des variations de l'épaisseur de l'aluminium et le l'oxyde sous le pad sont également démontrées, en revenant aux propriétés élastiques qui sont beaucoup moins coûteuses en terme de temps de simulation.

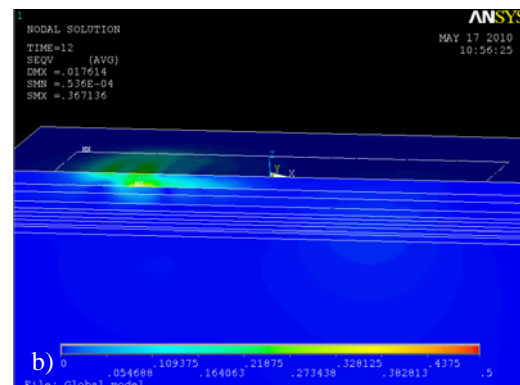
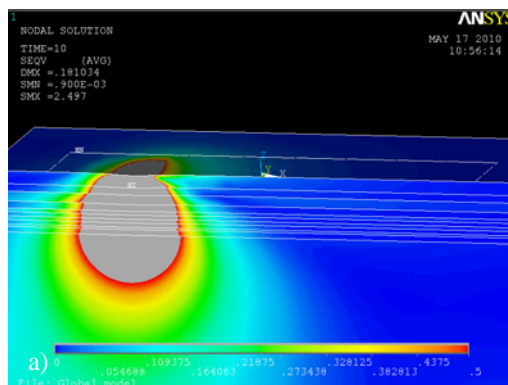


Figure 45. Contraintes Von Mises (a) et résiduelles (b) à la fin du premier passage.

Finalement, le modèle numérique 3D implicite est mis en place et s'est révélé capable de reproduire les divers paramètres de l'EWS et de l'architecture des pads.

V.B.I.c) Comparaison des deux modèles numériques

Dans ce paragraphe, les modèles explicite et implicite sont comparés, en se focalisant sur leur capacité à permettre d'étudier les différents paramètres de test et d'architecture des pads. Les modèles sont relativement complémentaires puisque le premier permet une meilleure étude des paramètres de test, mais moins bonne et beaucoup plus longue quand il s'agit de paramètres de dessins du pad.

		Modèle 3D explicite	Modèle 3D implicite
Paramètres de test EWS	Dynamique de la pointe	Très efficace	Non disponible
	Vibrations et vitesse du plateau portant le wafer	Très efficace	Non disponible
	Déplacement vertical	Très efficace	Très efficace
	Nombre de passages	Modéré, demande un long temps de calcul, possible uniquement via les forces de réactions	Modéré, via l'utilisation de propriétés élasto-plastiques, pas de représentation du contact
	Autres paramètres de test (attentes générales)	Efficace	Modéré, lié aux conditions de chargement et à l'absence du contact
Architectures de pad	Effet des épaisseurs de couche (aluminium, oxyde...)	Très efficace	Très efficace
	Effet de design (ex. présence de vias)	Bons résultats mais modéré dû à l'utilisation des propriétés homogénéisées	Très efficace
	Agencement du pad (ex. lignes vs. slots)	Modéré : même remarque	Très efficace
	Autres paramètres d'agencement (attentes générales)	Modérés dû à la non-représentation des détails	Très efficace : Modèle développé spécifiquement pour les comparaisons d'architecture de pad

Tableau 3. Capacité des modèles numériques à reproduire les paramètres de test et de dessin.

VI. Conclusions

Au cours de cette thèse, les paramètres de test et d'assemblage situés au niveau du pad ont été étudiés. Ces processus s'incluent dans le schéma de réalisation, permettant de vérifier son bon fonctionnement (EWS) ou de relier la puce au boîtier (assemblage filaire). L'objectif final est de développer des méthodes visant à réduire le nombre de problèmes mécaniques au niveau des interconnexions. Dans un premier temps, des tests, accompagnés d'observations, sont réalisés afin d'avoir une meilleure compréhension de la localisation et des causes des défaillances. De là, les paramètres induisant les fissures principalement dans l'oxyde situé sous le pad sont identifiés. Dans un second temps, la nanoindentation, au moyen des pointes Berkovich et cube corner, est utilisée pour évaluer la résistance des pads face à la compression. En parallèle, un modèle numérique est développé et calibré. Une fois les paramètres de test et d'assemblages étudiés et les tendances déduites, différentes architectures de pads sont évaluées. En effet, les effets de l'épaisseur des couches d'aluminium et d'oxyde, de la densité de cuivre, la présence de via et l'agencement de la couche supérieure du cuivre sur la robustesse mécanique des pads sont évalués au moyen de ces tests réels, de la nanoindentation et de la simulation. Globalement, des tendances similaires sont obtenues avec les différents outils, ce qui permet d'avoir confiance en leur capacité à reproduire le test de tri des plaques et l'assemblage filaire. La thèse se déroulant dans un contexte industriel, la dernière partie a pour but de détailler les règles d'utilisation de la nanoindentation dans le cadre de l'étude des pads. D'autre part un modèle éléments finis est développé afin d'obtenir plus rapidement que le modèle précédent des tendances fiables pour évaluer les architectures

dans le cadre du développement technologique. Finalement, la nanoindentation et la modélisation numérique constituent des alternatives fiables et moins coûteuses aux tests et assemblages effectués au niveau des pads. De plus des règles de dessins sont démontrées dans l'optique de fabriquer des structures mécaniquement plus robustes.

Les perspectives de ces travaux concernent la mesure des contraintes résiduelles dans les interconnexions liées à la fabrication des puces, mais également durant les étapes de tests (EWS) et d'assemblage. D'autre part, de nombreuses questions se posent encore en ce qui concerne l'initiation et la propagation des fissures durant l'EWS et la connexion filaire, que ce soit au niveau du mécanisme de leur apparition ou au niveau des seuils critiques à atteindre.

Références bibliographiques

[Anderson, 1957] Anderson, O.L., Christensen, H., et Andreatch, P., Technique for Connecting Electrical Leads to Semiconductors, Journal of Applied Physics 28, No. 8 (1957), pp. 923.

[Ansys, rel. 11] Ansys user manual, release 11.

[Bucaille, 2003] Bucaille, J.L., Strauss, S., Felder, E., Michler, J., Determination of plastic properties of metals by instrumented indentation using different sharp indenters, Acta materialia, vol. 51, 2003, pp. 1663-1678.

[Cacho, 2007] Cacho, F., Vettori, R., Galbiati, G., Thiery, G., Fiori, V., Ainaoui, M., Delpech, P., Tavernier, C. et Jaouen, H., Study of probing test for electrical wafer sort application: dynamical simulation and experimental measurement, IEEE 9th Electronics packaging technology conference, 2007, pp.391-395.

[Doerner, 1986] Doerner, M.F. et Nix, W.D., A method for interpreting the data from depth-sensing indentation instruments, Journal of material research vol. 1 (4), Juillet/Août 1986, pp. 601-609.

[Fiori, 2005] Fiori, V. et Orain, S., A multi scale finite element methodology to evaluate wire bond pad architectures, IEEE International conference on thermal, mechanical and multi-physics simulation experiments in microelectronics and micro-systems, Eurosime, 2005, pp. 648-655.

[Harman, 1978] Harman, G.G. et Cannon, C.A., The microelectronic wire bond pull test – how to use it, how to abuse it, IEEE transactions on components, hybrids and manufacturing technology, vol. chmt-1, no. 3, Septembre 1978.

[Harman, 1997] Harman, G.G., Wire bonding in microelectronics: materials, processes, reliability and yield, deuxième édition, McGraw-Hill.

[Moore, 2000] Moore, T.M., Hartfield, C.D., Anthony, J.M., Ahlburn, B.T., Ho, P.S. et Miller, M.R., in: D. G. Seiler, A. C. Diebold, T. J. Shaffner, R. McDonald, W. M. Bullis, P. J. Smith, E. M. Secula (Eds.), Characterization and Metrology for ULSI Technology 2000, Gaithersburg, USA, 26-29 Juin, 2000, American Institute of Physics Symposium Proceedings 550, (2001), pp. 431-439.

[Oliver, 1992] Oliver, W.C. et Pharr, G.M., An improved technique for determining hardness and elastic modulus using load and displacement sensing indentation experiments, Journal of materials research 1992, vol. 7 (6), pp. 1564-1583.

[Qin, 2010] Qin, I., Shah, A., Huynh, C., DeAngelis, D., Meyer, M., Mayer, M. et Zhou, Y., Thermosonic Au ball bonding process investigated using microsensor and laser vibrometer, IEEE electronic components and technology conference, 2010, pp. 1776-1782.

[Roucou, 2010-A] Roucou, R., Fiori, V., Cacho, F., Inal, K. et Boddaert, X., Evaluation of probing process parameters and pad designs: experiments and modelling correlations for solving mechanical issues, IEEE International conference on thermal, mechanical and multi-physics simulation experiments in microelectronics and micro-systems, Eurosime, 2010, pp 1-6.

[Roucou 2010-B] Roucou, R., Fiori, V., Inal, K. et Jaouen, H., Mechanical issues induced by electrical wafer sort : correlations from actual tests, nanoindentation and 3D dynamic modeling, IEEE Electronic system integration technology conference, 2010, pp. 1-6.

[Shah, 2008] Shah, A., Mayer, M., Zhou, Y., Hong, S.J. et Moon, J.T., In situ ultrasonic force signals during low-temperature thermosonic copper wire bonding, Microelectronics engineering, vol. 85, issue 9, pp. 1851-1857.

[Strom, 1998] Strom, J., Multi tier probe cards and contact resistance, Presentation at the IEEE Semiconductor wafer test workshop, 1998, <http://www.swtest.org/archive.html>, consulted on August, 18, 2010.

[Uno, 2009] Uno, T., Terashima, S. et Yamada, T., Surface enhanced copper bonding wire for LSI, IEEE electronics components and technology conference, 2009, p. 1486-1495.

[Vallauri, 2007] Vallauri, R., Gervasoni, M., Crippa, R. et Lazzari, S., Route60: a new vertical probing technology, IEEE Semiconductor wafer test workshop 2007, 3-6 Juin.

[Vath, 2009] Vath III, C.J., Gunasekaran, M. et Malliah, R., Factors affecting the long term stability of Cu/Au bonds subjected to standard and extended HTS, IEEE Electronics packaging technology conference, 2009, pp. 374-380.

[Wu, 2006] Wu, F. et Gonchond, J.-P., Optimisation de la technique de nanoindentation pour la caractérisation des propriétés mécaniques des oxydes basse permittivité, Université Joseph Fourier, Grenoble, France, Rapport de stage Mastère, 2006.

[Xu, 2008] Xu, H., Liu, C., Silberschmidt, V.V. and Wang, H., Effect of process parameters on bondability in thermosonic copper ball bonding, IEEE Electronic components and technology conference, 2008, pp. 1424-1430.

[Zhong, 2007] Zhong, Z.W., Ho, H.M., Tan, Y.C., Tan, W.C., Goh, H.M., Toh, B.H. et Tan, J., Study of factors affecting the hardness of ball bonds in copper wire bonding, Microelectronics engineering, vol. 84, 2007, pp. 368-374.

N° d'ordre : 2010 EMSE 0589

Romuald ROUCOU

EXPERIMENTAL AND NUMERICAL ANALYSES OF THE LOCAL MECHANICAL
FAILURES IN THE INTERCONNECTIONS INDUCED BY PARAMETRIC TESTS AND
ASSEMBLIES: OPTIMIZATION OF THE PROCESSES AND PAD ARCHITECTURES

Speciality: Microelectronics

Keywords: electrical wafer sort, wire bonding, finite element model, nanoindentation, pad architecture, mechanical failures

Abstract:

The diminution of the critical dimensions in the semiconductor industry and the introduction of new brittle dielectric materials raise questions on the mechanical resistance of the die and the pad architectures. Nowadays, pad structures are prone to crack. More precisely, the electrical wafer sort (EWS), which allows checking the electrical functionality of the die, and the assemblies such as the wire bonding to achieve the electrical connections with the packaging, are performed at the wafer level and introduce high levels of local mechanical stresses. Indeed, during these operations, failures in the oxide layers of the interconnections are observed. Experimental techniques (e.g. optical profilometry, FIB/SEM) are developed after EWS and bonding with gold and copper wires to gain insight on the root causes and localization of the failures. Some designs of experiments are set up to evaluate the influence of the test and process parameters (e.g. through wire pull and shear tests performed after bonding) and also of the various pad designs on the mechanical robustness of the structures. Thus, among others, the protective effect of the aluminium layer and the layout of the copper are demonstrated to be able to strengthen the structure. In addition, a novel analysis procedure, based on nanoindentation technique, is employed to compare various pads, which are complex multilayer systems. Moreover, several finite element models, using both explicit and implicit schemes are developed to mimic the EWS test. Indeed, these models have shown their ability to reproduce the loading conditions, the contact between the testing needle tip and the pad, and some inertial effect during the test. Finally, a comprehensive set of tools to evaluate and optimize the pad architectures is presented. Guidelines for pad layouts are also given, providing integration insights in the frame of the technology development.

École Nationale Supérieure des Mines
de Saint-Étienne

N° d'ordre : 2010 EMSE 0589

Romuald ROUCOU

ANALYSE EXPERIMENTALE ET NUMERIQUE DES DEFAILLANCES MECANQUES
LOCALES INDUITES DANS LES INTERCONNEXIONS PAR LES TESTS
PARAMETRIQUES ET ASSEMBLAGES : OPTIMISATION DES PROCES ET DES
ARCHITECTURES DES PLOTS DE CONNEXION

Spécialité: Microélectronique

Mots clefs : test paramétrique, connexion filaire, modèle élément finis, nanoindentation, plot de connexion, défaillances mécaniques

Résumé :

La diminution des dimensions critiques dans l'industrie du semi-conducteur requiert l'utilisation de nouveaux matériaux fragiles qui dégradent la résistance mécanique des puces. On s'intéresse plus particulièrement aux étapes précédant la mise en boîtier, à savoir les tests paramétriques qui permettent de vérifier la fonctionnalité électrique de la puce, et les assemblages tels la connexion filaire qui ont pour but d'établir les connexions avec le boîtier. Durant ces opérations, des défaillances mécaniques sont observées dans les interconnexions situées sous le pad. Des techniques expérimentales (par ex : FIB/MEB) sont mises en œuvre une fois les tests ou les assemblages avec des fils d'or et de cuivre réalisés afin de mieux comprendre les raisons d'apparition de ces défaillances ainsi que leur localisation. Des plans d'expériences sont mis en place pour évaluer l'influence des divers paramètres de tests et d'assemblage et également celle des architectures de pad, pour connaître ceux qui induisent les fissures dans l'oxyde. Ainsi, sont démontrés entre autres l'effet protecteur de l'aluminium et l'influence du design des architectures. En parallèle, une nouvelle méthode d'analyse basée sur la nanoindentation est utilisée pour comparer la robustesse mécanique de divers plots de connexion. Il s'agit ici d'étudier un ensemble multicouche, ce qui pose de nombreux challenges au niveau de la fiabilité des mesures et des résultats. D'autre part, plusieurs modèles éléments finis complexes, prenant en compte la gestion du contact entre la pointe de test et le pad ainsi que les effets inertiels associés, sont développés dans le but de reproduire les conditions de chargement sur les pads. Finalement, un ensemble d'outils adaptés à l'étude et l'optimisation des architectures de pad, dans une optique industrielle, est présenté de même que des règles de dessin permettant d'accompagner le développement technologique.